DOCUMENT RESUME

ED 181 055 TM 009 959

TITLE	Proceedings of The Invitational Conference on Testing Problems (24th, New York, New York, October 29, 1960).
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	29 Oct 60
NOTE	138p.
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Administrator Evaluation: Conference Reports: Developing Nations: Educational Problems: *Educational Research: Elementary Secondary Education: Federal Aid: Federal Programs: Foreign Countries: *Language Tests: Occupational Tests: Prose: *Talent Development: Talented Students: Teacher Behavior: *Testing: *Testing Problems
IDENTIFIERS	Africa: *Cooperative Research Program: In Basket Simulation

ABSTRACT

This conference focused on two themes: research in testing and the cooperative research program, and testing in the language arts. The morning session was concerned with the impact of the federal Cooperative Research Program on educational research. Papers were entitled: The Support of Measurement Projects by the Cooperative Research Program. (Hall, Hjelm, and Conrad): In-Basket Tests and Factors in Administrative Performance (Frederiksen): Models of Teacher Behavior in the Classroom (Travers): and the Prediction of Talented Behavior in the Junior High School (McGuire). Papers for the afternoon session were entitled: Testing for Elegance (Martin): Competency First: New Tests in Foreign Languages (Starr): and Estimating Structure in Prose (Lorge). The luncheon address, The Pace of Change, was delivered by Arthur S. Adams and recounted his experiences as an educational consultant in emerging African nations. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED181055

INVITATIONAL CONFERENCE ON TESTING PROBLEMS

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

1960

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Educational
Testing Service
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

TM009959

Proceedings

EDUCATIONAL TESTING SERVICE

Copyright, 1961, Educational Testing Service

Princeton, N. J.

Library of Congress Catalog Number: 47-11220

Printed in the United States of America

Invitational Conference on Testing Problems

October 29, 1960
Hotel Roosevelt
New York City

JOHN B. CARROLL
Chairman

Educational Testing Service
Princeton, New Jersey
Los Angeles, California

ETS Board of Trustees 1960-61

James B. Conant, *Chairman*

Arthur S. Adams
Frank D. Ashburn
George F. Baughman
William G. Carr
John W. Gardner
Caryl P. Haskins
Stuart Hedden
Philip J. Hickey
Arnold E. Joyal
Katharine E. McBride
John D. Millett
Edward S. Noyes
B. Alden Thresher
O. Meredith Wilson

ETS Officers

Henry Chauncey, *President*
William W. Turnbull, *Executive Vice President*
Henry S. Dyer, *Vice President for College Board Programs*
Robert L. Ebel, *Vice President for General Programs*
G. Dykeman Sterling,
Vice President for Administration and Treasurer
Catherine G. Sharp, *Secretary*
Paul G. Haskell, *Assistant Treasurer*

Foreword

THERE were two themes upon which the 1960 Invitational Conference on Testing Problems focused its attention: "Research in Testing and the Cooperative Research Program" and "Testing in the Language Arts." The speakers covered a wide range of interests within both these areas. Their papers and the discussion of them that followed provided a program that was pertinent and provocative—valuable to those who specialize in the problems of measurement and stimulating to all who are concerned about the problems of education in general.

Dr. John B. Carroll of Harvard University who, as Conference Chairman, was responsible for assembling the program and speakers, deserves our praise and appreciation. I would also like to express our thanks to the distinguished speakers whose papers we are pleased to publish in this edition of the *Proceedings*.

Henry Chauncey

President

Preface

TWO IDEAS entered my mind as I cast about for themes around which to build the program for the 1960 Invitational Conference on Testing Problems. One came from the fact that the interest of the Federal Government in educational research has in the last few years increased to a point where it is a major source of support for programs of research in education, including research in educational tests and measurements. The other was my own special interest in problems of testing in the "language arts." Both these ideas seemed to have wide enough appeal to serve as themes for the Conference, and they had interesting connections with each other: it happened, for example, that several developments in language testing had been made possible as a result of government support.

As it turned out, this interplay of ideas made for a most interesting and stimulating conference, when brought to reality by a panel of speakers who were not only well informed but also able to make creative and thought-provoking contributions. It is a measure of the success of the Invitational Conferences that so many people—this year more than 550—make the effort to come to the Hotel Roosevelt to hear the speakers in person and listen to the sometimes heated discussion. I cannot believe that many were disappointed this year.

The morning session was concerned with the impact of just one major program of Federal support to educational research—the Cooperative Research Program administered by the U. S. Office of Education. The first paper, read by Herbert S. Conrad for Roy M. Hall (unfortunately absent because of illness), was a description of the program in general. The remaining three papers were by the directors of three out of the many research projects supported by this program: Norman Frederiksen, Educational Testing Service; Robert M. W. Travers, University of Utah; and Carson McGuire, University of Texas.

Adjourning for lunch, the conferees had the pleasure of hearing Dr. Arthur Adams, President of the American Council on Education, brilliantly and entertainingly speak on "The Pace of Change." Dr. Adams described some of his experiences as an educational consultant in the new African nations and reminded us of the pace at which changes are occurring in all phases of education, international as well as domestic, and of our responsibilities in keeping abreast of them.

The afternoon session turned to the theme of "Testing in the Language Arts." The speakers were Harold C. Martin, Harvard University; Wilmarth H. Starr, Modern Language Association and New York University; and Irving Lorge, Teachers College, Columbia University.

I am delighted to have this opportunity to express my great appreciation to the speakers for the thought and effort they devoted to the preparation of their remarks, and for the excellent presentations they made at the Conference.

It is also a pleasure to have my turn at representing the profession at large, and the Conference invitees in particular, in thanking the Educational Testing Service—its President, Henry Chauncey, and all its staff members who contributed to the success of this annual event—for providing so happy an occasion at which those concerned with problems of educational measurement can meet together and exchange information and ideas.

John B. Carroll
Chairman

CONTENTS

3 Foreword by HENRY CHAUNCEY
5 Preface by JOHN B. CARROLL

Session I

9 Remarks of the Chairman, **John B. Carroll**, Professor of Education, Graduate School of Education, Harvard University
12 The Support of Measurement Projects by the Cooperative Research Program, **Roy M. Hall**, **Howard F. Hjelm**, **Herbert S. Conrad**, Cooperative Research Program, U.S. Office of Education
21 In-Basket Tests and Factors in Administrative Performance, **Norman Frederiksen**, Research Division, Educational Testing Service
38 Models of Teacher Behavior in the Classroom, **Robert M. W. Travers**, Department of Educational Psychology, University of Utah
46 The Prediction of Talented Behavior in the Junior High School, **Carson McGuire**, Department of Educational Psychology, University of Texas

Luncheon Address

74 The Pace of Change, **Arthur S. Adams**, President, American Council on Education

CONTENTS Continued

Session II

85 Remarks of the Chairman

87 Testing for Elegance, **Harold C. Martin**, General Education A, Harvard University

97 Competency First: New Tests in Foreign Languages, **Wilmarth H. Starr**, Foreign Language Testing Program, Modern Language Association of America

111 Estimating Structure in Prose, **Irving Lorge**, Institute of Psychological Research, Teachers College, Columbia University

123 List of Conference Participants

Session I

Remarks of the Chairman

JOHN B. CARROLL,
*Roy E. Larsen Professor of Education,
Harvard University*

OUR THEME this morning is "Research in Testing and the Cooperative Research Program." It is prompted by the recently growing and massive interest in educational research on the part of the Federal Government. To be sure, the government has been involved in educational research, of a sort, for a good number of years. Group testing received its first impetus from the Army program of World War I, and the programs of military personnel research engendered by World War II are well known to all. But most of this was only on military problems; any relevance to educational problems in the schools was incidental. After many years in which the Federal Government paid scant attention to educational research and testing in the public schools, the Congress suddenly had a change of heart and in 1954 passed Public Law 531, which gave rise to the Cooperative Research Program of the United States Office of Education.

It is heartening to note that this program was not one of the reactions to the orbiting of the first Sputnik. Much of the impetus for the program came from the interest of legislators in the problems of mental retardation. Having looked recently into the research literature on mental retardation, I can attest that there was ample reason to put funds into this kind of research. But however costly the care of mental defectives is for our society, these people constitute only perhaps two or three per cent of the population of school-going age; setting up a Cooperative Research Program with more than half of the funds originally allocated to research in mental retardation was a case of letting the tail wag the dog. But at any rate let us be thankful for the dog, with or without tail.

Then came Sputnik I and another wrong reason (however happy the result) for increasing Federal aid to educational research. The outcome was the National Defense Education Act of 1958. Here again the law exhibited what many have regarded as an improper balance among the various kinds of support that could have been given. The emphasis was upon science, mathematics, and foreign languages, with a substantial bow in the direction of school guidance programs and the development of educational "media" such as television, films, and even teaching machines. By the quirks of the legislative process, educational research under the National Defense Education Act has been limited to research in foreign language teaching and in the development of educational media. Even research in the teaching of science and mathematics was excluded unless it could be brought under the heading of media research. Nevertheless, we can be thankful for the many good things brought about by the NDEA.

It seemed therefore fitting to dedicate a part of our program today to a recognition of these flourishing programs of Federal support to educational research and testing. Through what we say here today we can signalize the importance we attach to educational research and its potential benefits to education. People professionally concerned with educational research and measurement are gratified to note the establishment of these programs of Federal support. This conference provides an occasion for looking at some of the first fruits of these programs.

I have had to be selective. This morning we will limit our consideration to the Cooperative Research Program of the United States Office of Education. Research and development programs supported by the National Defense Education Act will come into our view only in connection with one of this afternoon's topics.

Our first speaker was to have been Dr. Roy M. Hall, Assistant Commissioner of Education for Research, and administrative head of the Cooperative Research Program. Dr. Hall is well known for his research and teaching in the field of educational administration, and was at the University of Texas before taking up his duties in Washington several years ago. Because of his temporary illness, Dr. Hall's paper will be read by Dr. Herbert S. Conrad, his associate in the U. S. Office of Education and a person whose work in educational measurement and statistics is familiar to you. Actually, Dr. Conrad was a co-author of the paper we are to hear, entitled, "The Support of Measurement Projects by the Cooperative Research Program."

Following this, you will hear three papers, each one by a director of one of the scores of projects which have been or are being supported by the Cooperative Research Program. I was faced with an embarrassment of riches as I tried to select interesting and provocative speakers; I shall not try to indicate the reasons for my choices—let us simply leave it that I attempted to bring you addresses which would give a satisfactory impression of the diversity and scope of the research being supported by the Cooperative Research Program.

The first of these papers is "In-Basket Tests and Factors in Administrative Performance," by Dr. Norman Frederiksen of the Educational Testing Service, describing results of a project on testing the administrative abilities of elementary school principals. For this project, ETS joined forces with Columbia University.

The second is "Models of Teacher Behavior in the Classroom," by Dr. Robert M. W. Travers, Chairman of the Department of Educational Psychology, University of Utah, and author of several books on measurement and educational research. This paper is in the nature of a progress report on his investigation of the process of teaching.

The third is "The Prediction of Talented Behavior in the Junior High School," by Dr. Carson McGuire, of the Department of Educational Psychology at the University of Texas, a psychologist who has long been interested in ways of measuring the less accessible aspects of human behavior. I am sure that his approach to the increasingly popular "search for talent" will be stimulating.

Roy M. Hall,
Howard F. Hjelm,
and Herbert S.
Conrad, *U. S. Office
of Education*

The Support of Measurement Projects by the Cooperative Research Program of the U. S. Office of Education

THE UNITED STATES OFFICE of Education provides support, through its Cooperative Research Program, for research of significance to education. This Program is both *extra-mural* and *cooperative*: "extra-mural," in that the research originates with, and is conducted by, investigators and institutions outside the Government; and "cooperative," in that those who originate the research share with the Office of Education a vital interest in the project and are willing to contribute toward its financing. A broadly representative group of educational researchers assists the Office in evaluating the proposed research.

Public Law 531, under which the Cooperative Research Program operates, authorizes the Office of Education to "enter into contracts or jointly financed cooperative arrangements with universities and colleges and State educational agencies for the conduct of research, surveys, and demonstrations in the field of education." The particular aspects of education eligible for support under the Cooperative Research Program are not specifically delineated in the law. Any project of significance to education is eligible for consideration. It may be a project concerned with pre-school, elementary, secondary, higher, or adult education. It may be concerned with public or private education. It may deal with the teaching-learning situation, administrative functions, sociological aspects of the classroom and community, psychological theory, or principles of measurement. It may deal with the means employed or the products achieved. It may deal with the past or the present, or, indeed, with the future. In essence, of course, all good research is future-minded: "the past is prologue," and by definition, the present doesn't last very long.

Research and development in the area of measurement deserve the interest and support of all educators: partly because measurement renders immediate practical service; but more important, because advances in measurement frequently spark significant advances in education. Many of the investigations supported by the Cooperative Research Program of the Office of Education involve the construction and refinement of measuring instruments. Some examples of such investigations—among our proudest products—will be presented by the other speakers on this morning's panel. In the time at my disposal, let me describe to you a variety of other studies under the Cooperative Research Program which deal with measurement.

Measurement Projects Supported by the Cooperative Research Program

One type of study supported by the Cooperative Research Program seeks to *refine existing instruments* (published or unpublished), *or to improve our understanding of them.* An example is the project just started by Dr. Wallace A. Kennedy, Chairman of the Human Development Clinic of Florida State University. Dr. Kennedy is standardizing the Third Revision of the Stanford-Binet Intelligence Scale on Negro elementary school children in the Southeastern United States. The norming sample will consist of 1,800 Negro students, grades 1 through 6, from five States in the Southeast. In order to assure representativeness, the sample will be stratified by such factors as urbanization, socioeconomic status, and educational facilities available. Biographic data will be collected in order to compare the representativeness of the sample with data from the United States Census Bureau. Dr. Kennedy plans to perform some validation analyses by correlating Stanford-Binet scores with achievement test scores and with academic performance. Item analyses, including factor analysis of item intercorrelations, will also be carried out. This study is scheduled for completion in about a year (December 1961).

A second type of study supported by the Cooperative Research Program includes the *development of new instruments*—often as part of a larger undertaking in the field of measurement. In this connection, the word "new" may require some qualification. Some tests are "newer" than others; it is difficult, sometimes, to distinguish between innovation, refinement, and recombination: all three may, of course, occur together.

Perhaps the best-known of the measurement projects supported by the Cooperative Research Program (in conjunction with other Government agencies) is the so-called "Project Talent." This study will survey the aptitudes, abilities, interests, and other characteristics of close to half a million pupils of grades 9, 10, 11, and 12 in approximately 1,000 high schools in the United States (both public and private). The study is directed by Drs. John C. Flanagan and John T. Dailey, under the auspices of the University of Pittsburgh. In order to assure representativeness of the sample, schools were stratified (with the aid of a modern, high-powered computer) on such factors as geographic location, size of the twelfth grade class, and holding power of the school. Follow-ups of students in the sample are planned at intervals of one, five, ten, and twenty years after graduation from high school.

The major goals of Project Talent, other than the development and norming of the various test-instruments, can be very briefly summarized as follows:

1. To provide a scientifically planned national inventory of youth.

2. To determine the specific patterns of aptitudes, abilities, and interests which are best adapted to various college courses and careers.

3. To carry out follow-up studies to determine the educational experiences which best contribute to the development of these aptitudes and abilities.

4. To determine the guidance procedures most effective in assisting each student to select a career which will assure him personal satisfaction and success.

The battery of data-collection instruments was developed especially for Project Talent. No commercial tests are included in the battery. There is a Student Information Blank, a Student Interest Inventory, and a Student Activity Inventory, plus three questionnaires to be completed by school personnel concerning the schools themselves. Numerous aptitude and achievement tests were administered to the students.

The Talent battery of tests is not, and will not become, available commercially. A limited number of sets of the battery can be made available, however, for qualified research.

Dr. Sarnoff A. Mednick of the University of Michigan has developed an associative theory of creative thinking, and has carried out some preliminary experiments which seem to support the theory. An objective test entitled Remote Associates Test, which is an operational statement of the theory, has been developed for measuring individual differences in creative talent.

Items in the Remote Associates Test present three words for which the subject is to respond with a fourth word which serves as a connective link among them. For example, three words such as the following are given:

surprise line birthday

The answer is party—surprise party, party line, birthday party.

Another example is

base snow dance

The correct response is ball—baseball, snowball, dance-ball.

Try this one.

railroad girl class

The Remote Associates Test is being administered experimentally in several graduate schools as a potential selection device. It appears to be a promising instrument, correlating in the low 30's with tests of verbal ability. Mednick's work has been conducted mainly at the college level. He has a Cooperative Research contract to extend the research to the secondary level; from this should develop several forms of the Remote Associates Test for high school students.

Frederick B. Davis and Gerald S. Lesser of Hunter College are coinvestigators on two Cooperative Research projects in which they have developed the Hunter Aptitude Scales for Gifted Children, to be used in identifying pre-school and elementary school children gifted in one or more of five areas. The five scales which were developed are as follows:

1. Space Conceptualization: a cluster of skills in judging spatial relationships, the sizes of objects, and movements in space.

2. Vocabulary: defined as memory for word meanings—

with reasoning elements, such as those required in a verbal analogies test, reduced to the minimum.

3. Number Ability: closely defined as memory of the fundamental combinations in addition, subtraction, multiplication, and division, as well as skill in using these accurately.

4. Reasoning: the ability to formulate concepts, to weave together ideas and concepts, and to draw inferences from them.

5. Science: knowledge of scientific principles and information.

These tests are all individually administered, and require no reading on the part of the examinees.

Working at the other end of the human-ability scale, Dr. Lloyd M. Dunn, Coordinator of Special Education, George Peabody College for Teachers, has developed The Peabody Picture Vocabularly Test (PPVT). The PPVT, which is being distributed by a test publisher, is designed to measure comprehension of the spoken word, including the ability to associate verbal symbols with pictorial representations of objects, events, and actions. The test consists of 150 items and has two forms. It is individually administered, yielding three types of derived scores—age scores, quotients, and percentiles. Although the PPVT was constructed primarily for use with mentally retarded individuals, it can be given to any English-speaking person between 2½ and 18 years of age who is able to hear words, see the drawings, and has the ability to indicate yes or no in a manner which communicates.

Drs. George G. Stern and Joseph M. Masling, Syracuse University, conducted a Cooperative Research Project entitled "Unconscious Factors in Career Motivation for Teaching." The purpose of this study was to develop suitable instruments for assessing unconscious factors in career motivation among teachers, and to relate such factors to performance in specific areas of the elementary school teaching process. The test, developed with the help of public school administrators, supervisors, and teachers, consists of descriptive sketches of a number of motivational types encountered among teachers. A sample of individuals considered representative of each type was then selected, interviewed, and tested. The data obtained from these subjects provided a basis for building objective diagnostic instruments for the identification of ten motivational types. These instruments were administered to samples of teacher-trainees and practice teachers. Classroom rating scales were also developed and tested in observations of elementary school classes.

An additional type of Cooperative Research investigation falls in the area of *application of measuring instruments to the educational program*. This involves some of the guidance and placement functions of the schools. One such study, "Investigation of the Use of Statistics in Counseling Students," was carried out by Robert E. Hewes, Registrar, and Richard W. Willard, Statistical Analyst, of the Massachusetts Institute of Technology. They investigated the potential usefulness of certain statistics as an aid to M.I.T. students faced with choosing their college careers. Profiles of the students on academic measures were compared with profiles of students in various majors. The profile analysis of the academic measures added no useful information regarding the students' predicted likelihood of graduation in the various majors. A student with a poor prognosis in a more difficult major had little better prognosis in a less difficult major. One may infer that, among the highly selected students composing the M.I.T. enrollment, differences must be accounted for in other terms, such as those of interest and motivation.

The Cooperative Research Program has also sponsored research investigations concerned with *increasing our basic knowledge of measurement theory*.

Dr. Julian C. Stanley at the University of Wisconsin is currently conducting a Cooperative Research Project entitled "Development and Analysis of Experimental Designs for Ratings." Studies involving ratings are usually partially invalidated by excessively high intercorrelations among the traits rated, since raters tend to rate individuals too much in terms of a single factor, such as scholastic ability or general attractiveness. Dr. Stanley plans to analyze completely a four-dimensional matrix, where each rater rates each individual more than once on each trait, for sources of bias. He hopes to provide a general analytical basis for minimizing biases and maximizing independence of traits. Statistical adjustments will be developed for use with incomplete designs in which not every rater rates every individual on every trait.

Another Cooperative Research investigation which may be of interest to you was a study conducted by Drs. Raymond C. Norris and Howard F. Hjelm at George Peabody College for Teachers, on the effects of non-normality upon the sampling distribution of the product-moment correlation coefficient. It was an empirical investigation, employing a high-speed electronic computer. Drs. Norris and Hjelm used experimental populations which were rectangular, leptokurtic, and skewed, having approximately no correlation and substantial correlation. Sampling distributions of the correlation coefficient were established for samples of 15, 30, and 90 cases. In general, the findings were that the effects of non-normality were not serious for populations with a true *r* close to zero; but were significant for populations having substantial correlation, with the nature and direction of the deviation specific to the type of non-normality in the population.

Measurement Projects for the Future

It is clear that the Cooperative Research Program has supported a variety of research investigations in the area of measurement. We would, however, like to support a still larger variety. Educational measurement is, to be sure, progressing: but in our judgment it is not progressing fast enough, on enough fronts, with enough coordination.

Lest that much-abused word, "coordination," cause alarm, let me express at once our profound conviction that coordination is far outweighed by inspiration. The kind of coordination we mean to stress is coordination between the experts in measurement and the persons and functions that measurement should serve. Experts tend to develop a "closed society," reveling in mutual appreciation and in the pursuit of happiness through the achievement of technical perfection—at the cost, sometimes, of appreciating others' needs. Closer liaison is needed between those who have ideas about curriculum, methods, supervision, school construction, school finance, public relations, delinquency, guidance, and so on, and those who have skills and ideas in the field of measurement. Too much of educational measurement has centered on the school child or college student; or, rather, not enough has been devoted to other areas and other persons in the whole field of education. Measurement should originate out of, and provide tools for, *all* the various aspects of education—from graduate school down to pre-school, from classroom teaching up to the schoolboard room and the legislature. This, as we see it, is the main challenge to educational measurement today.

There are some other challenges, too—all, as we see it, growing out of the values, standards, or (if you will) folkways of American society. For one thing, we honor the dignity of man and respect the individual: so the standards of measurement are geared, preferably, to the individual. Yet education, applied to nearly 50,000,000 persons, must be concerned with *groups;* and in some delicate domains, where the pressure for conformity of response may exceed the impulse toward truth, group technics of measurement may add to validity. Another characteristic of our society seems to be its preoccupation with the present, or, at least, its limited concern for the future—and so installment debt multiplies; overcrowding of schools is corrected after the event, but not prevented; etc. In the field of measurement, our standards seem to be fixed largely in the framework of the *present.* Actually, however, school children grow and develop; what we need are devices accurate enough to measure not only present status, but also *change* or *difference.* Much more attention is needed to this requirement. For this purpose, of course, tests far more accurate than those useful for group measurement are essential; in fact, we currently have very few tests sufficiently accurate to measure reliably the growth and development of individual school children. A somewhat related problem which is not receiving the intensity of research that its importance justifies is the problem of differential prediction. Here work is needed both on the tests (or other devices) and on the criteria. Finally, Americans are generally accused of being charmed to excess by considerations of speed, economy, and objectivity—and so we have the whole field of measurement dominated by the paper-and-pencil, machine-scorable format. It seems fair to ask whether the virtues of speed, economy, and objectivity may have been gained at the expense of innovation, and scope, and validity.

The Cooperative Research Program is open to investigations along all the lines indicated above—as well, of course, as others. We do not pretend to anticipate the ideas of able investigators.

Concluding Remarks

In closing, let me say that the Cooperative Research Program of the U. S. Office of Education is anxious to support as many research projects in education as it can. If you are planning some research of significance to education; if the research design is sound; if the personnel and facilities are adequate; if the cost is justifiable; and if you need financial assistance, you should consider sending your proposed plan of research to the Cooperative Research Program. We invite and encourage you to submit applications for Cooperative Research funds to conduct research concerned with measurement.

In-Basket Tests and Factors in Administrative Performance

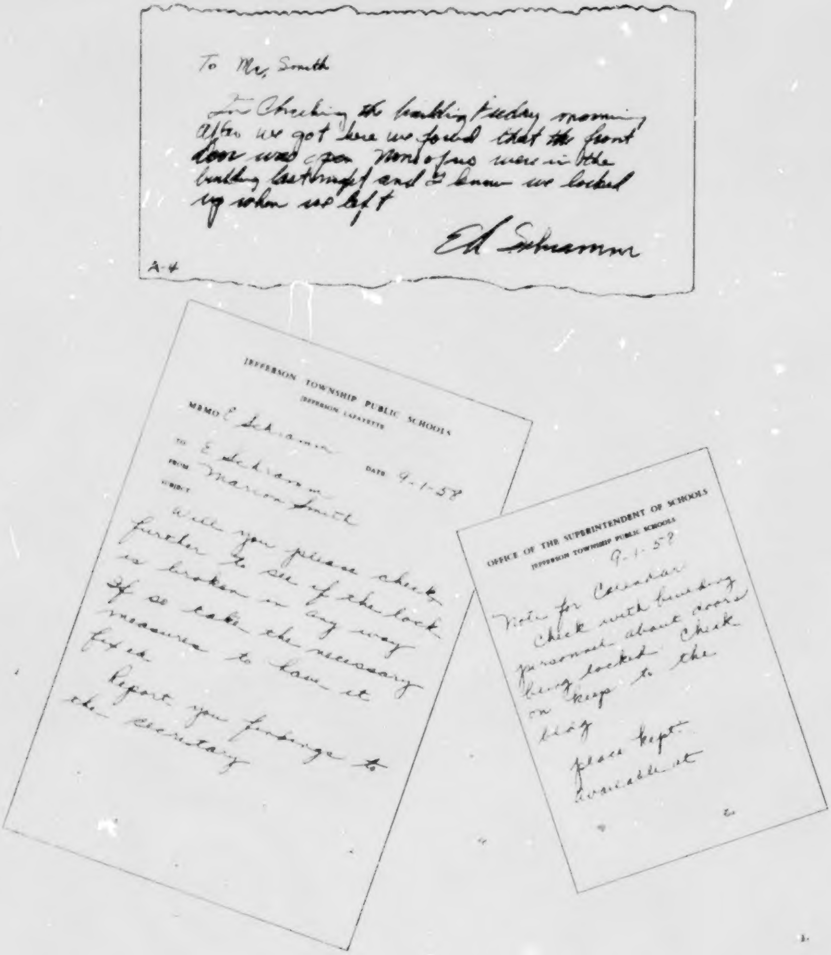
NORMAN FREDERIKSEN,
*Research Division,
Educational Testing
Service*

THE PURPOSE of the study which I will describe this morning was to learn more about performance in educational administration. There is a tendency, in evaluating administrators, to think in terms of one over-all dimension of goodness or badness. We felt that formulating the evaluation problem in terms of one variable was too simple. One of our purposes was to try to identify some major dimensions of administrative behavior; another was to discover how such dimensions might be related to a variety of other measurable characteristics of people.

The study was a cooperative one, done by staff members of Teachers College, Columbia University, and Educational Testing Service. When I say *we* I am referring to Dr. Daniel Griffiths of Teachers College, Director of the Project; Dr. John K. Hemphill of ETS, Associate Director; Dr. Glen Stice, Research Associate at ETS, and myself. The study was supported by a grant to Teachers College from the Cooperative Research Program of the U. S. Office of Education.

One difficulty encountered when one tries to study behavior on the job is that it is impossible to tell whether variation in performance should be attributed to differences among the subjects or to differences among the jobs. In this study we eliminated the job as a source of variation by simulating a school and community and giving each one of the 232 subjects the job of being principal of the same simulated school. Enough information about this school and community was provided that subjects could reasonably be expected to take action on the administrative problems presented to them.

The subjects of the investigation were 232 elementary school principals. They came from school districts all over the



To Mr, Smith

In Checking the building tuesday morning After we got here we found that the front door was open None of us were in the building last night and I know we locked up when we left

Ed Schramm

A-4

JEFFERSON TOWNSHIP PUBLIC SCHOOLS
JEFFERSON LAFAYETTE

MEMO E Schramm

TO E Schramm DATE 9-1-58

FROM Marion Smith

SUBJECT

Will you please check further to see if the lock is broken in any way If so take the necessary measures to have it fixed.

Report your findings to the secretary

OFFICE OF THE SUPERINTENDENT OF SCHOOLS
JEFFERSON TOWNSHIP PUBLIC SCHOOLS

9-1-58

Note for Calendar
Check with building personnel about doors being locked. Check on keys to the bldg
place kept available at

Exhibit 1

United States, and they constituted quite a varied group in terms of age, experience, and ability.

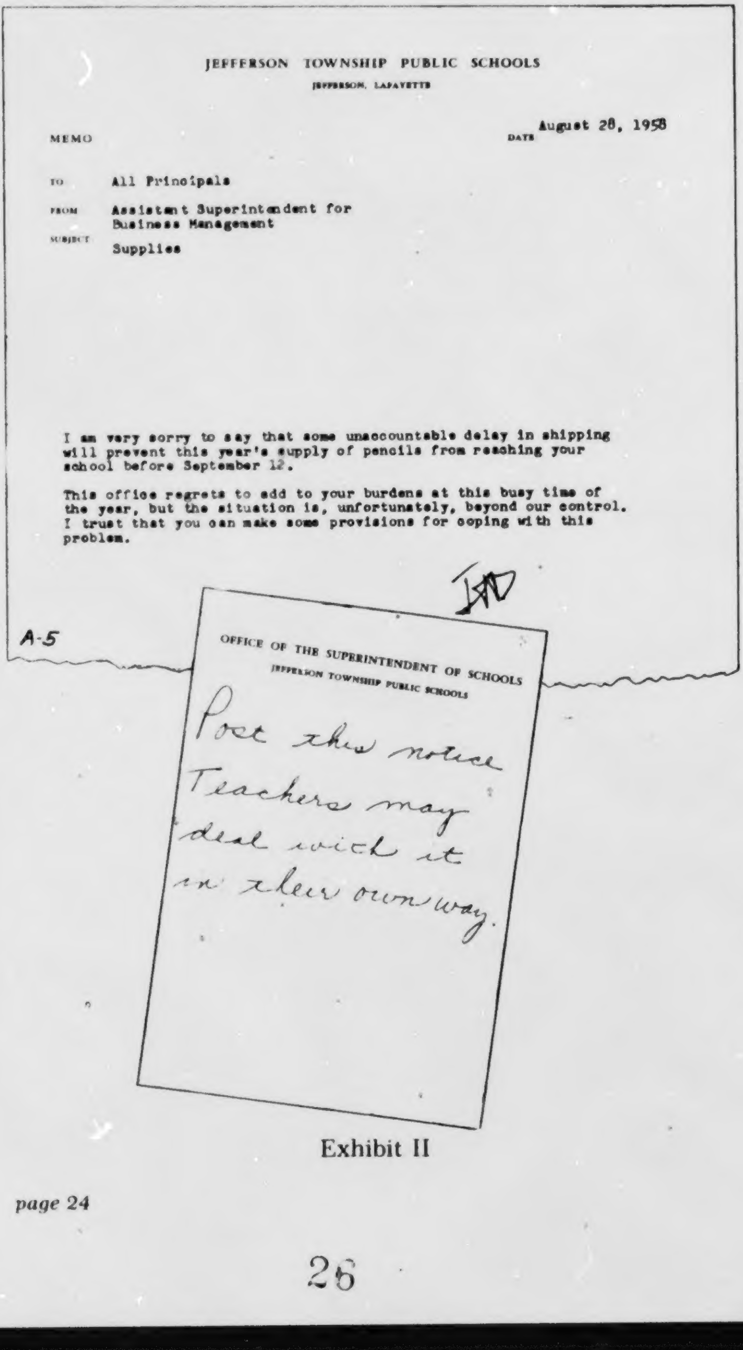
The test is a fairly elaborate one, requiring five days to administer. About 20 people were tested at one time. The test began on Monday morning, when the examinees were introduced to the school of which each was to be the new principal. The simulated school is Whitman School, located in Jefferson in the hypothetical state of Lafayette. Each subject was temporarily given a new name, Marion Smith. The subjects were instructed that they were not to play a role, that each was to bring to the new job his own background of experience, his own knowledge and personality. During the test he was to perform the duties of the new principal of Whitman School.

The participants spent the first day and a half in learning about Whitman School and its community. First, a film strip was presented which gave the participants an over-all view of Jefferson. Then the principals were given an opportunity to study the Jefferson School-Community Survey which had recently been completed by the School of Education at Lafayette State. Next the group viewed a sound color film which took the subjects inside Whitman, where they saw the faculty and children at work. The subjects were given personnel folders for the teachers and staff, a floor plan of the school, and a staff roster. Study guides were provided to direct the subjects' study of the materials.

On Tuesday morning the indoctrination was continued with additional material for study, including a Staff Handbook, the School Board Handbook, excerpts from the Lafayette school law, copies of the most recent school census, a class size list, the school calendar, and a report of achievement test scores. All this printed material became the principals' files throughout the week. The examinees also listened to tape recordings of school board meetings and conversations involving teachers and parents.

At the end of the day and a half of orientation, the subjects had as much information as would be expected of a new principal in an actual situation. It was now reasonable to expect the subjects to take action on problems arising in the administration of the school.

The balance of the week was devoted to work sessions in



JEFFERSON TOWNSHIP PUBLIC SCHOOLS
JEFFERSON, LAFAYETTE

MEMO

DATE August 28, 1958

TO All Principals

FROM Assistant Superintendent for Business Management

SUBJECT Supplies

I am very sorry to say that some unaccountable delay in shipping will prevent this year's supply of pencils from reaching your school before September 12.

This office regrets to add to your burdens at this busy time of the year, but the situation is, unfortunately, beyond our control. I trust that you can make some provisions for coping with this problem.

A-5

OFFICE OF THE SUPERINTENDENT OF SCHOOLS
JEFFERSON TOWNSHIP PUBLIC SCHOOLS

Post this notice
Teachers may
deal with it
in their own way.

Exhibit II

which each Marion Smith performed the duties of the principal of Whitman School. All the participants were presented with the same set of administrative problems under the same conditions. Each new principal was given memo pads, letterheads, paper, pencils, and paper clips. He was instructed merely to be the principal. He was not to say what he *would* do—he was to *do* it. He actually wrote memos, called meetings, prepared agendas, made notes in preparation for interviews, and the like.

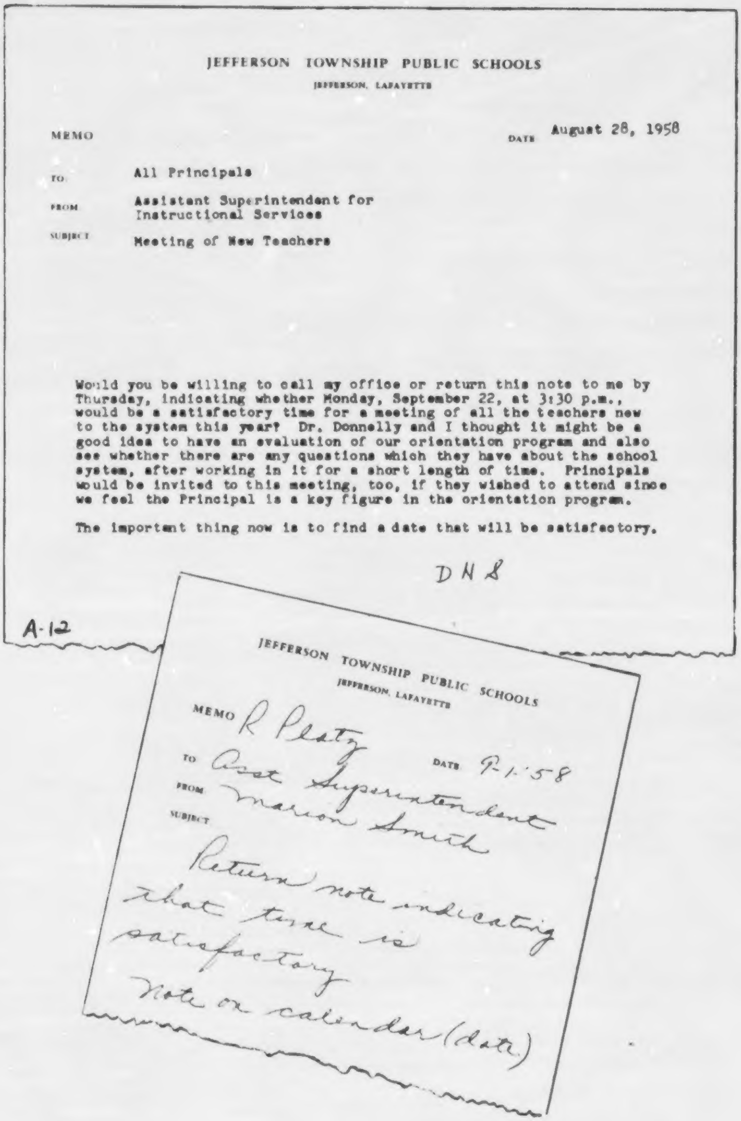
The primary method of presenting problems in our study was the in-basket test. Such a test consists of facsimiles of the letters, memoranda, and other contents of the in-basket such as is found on every administrator's desk. Exhibits 1, 2, and 3 in your handout show some in-basket items and the responses to them made by one of our subjects. Four in-basket tests were presented, each requiring a half-day session. About thirty-two problems were included in each in-basket; they were chosen in the light of a theoretical formulation of the job of an elementary school principal.

In addition to the desk work involved in the four in-baskets, the subjects were required to participate in committee work, observe the work of teachers shown in kinescopes, and react to conference situations presented by means of tape recordings.

This is a very brief description of a rather elaborate situational test. The participants tell us that the situation was realistic and that they built up vivid images of the people involved. Each participant received the same opportunities to learn the background, and each was presented with the same problems under identical conditions. Therefore we can attribute differences in behavior to the participants rather than to variations among jobs.

Each principal left an envelope full of memos, letters, reminders, instructions to his secretary, appointment calendars, and the like. How can this material be scored?

The first step towards developing a scoring procedure was to examine in-basket responses to see how respondents differ. As a result of such examination by a number of observers, a large pack of cards was collected, each card containing a statement of some kind of difference observed in the way the



JEFFERSON TOWNSHIP PUBLIC SCHOOLS
JEFFERSON, LAFAYETTE

MEMO

DATE August 28, 1958

TO: All Principals

FROM Assistant Superintendent for Instructional Services

SUBJECT Meeting of New Teachers

Would you be willing to call my office or return this note to me by Thursday, indicating whether Monday, September 22, at 3:30 p.m., would be a satisfactory time for a meeting of all the teachers new to the system this year? Dr. Donnelly and I thought it might be a good idea to have an evaluation of our orientation program and also see whether there are any questions which they have about the school system, after working in it for a short length of time. Principals would be invited to this meeting, too, if they wished to attend since we feel the Principal is a key figure in the orientation program.

The important thing now is to find a date that will be satisfactory.

DNS

A-12

JEFFERSON TOWNSHIP PUBLIC SCHOOLS
JEFFERSON, LAFAYETTE

MEMO R Pleatz

TO Asst Superintendent

DATE 9-1-58

FROM Marion Smith

SUBJECT

Return note indicating that time is satisfactory

Note on calendar (date)

Exhibit III

principals behaved. The cards contained phrases like "compulsive," "postpones decisions," and "makes unwarranted assumptions." A second source of ideas for scoring categories came from theories of administration and leadership, especially the theories of the Director and Associate Director of the project.

Eventually all these modes of behavior were collapsed to make the 68 scoring categories shown in Table 1 of your handout. A scoring manual was written which gives appropriate definitions and rules for scoring. Each in-basket problem can be thought of as an opportunity to display the behavior described by the category. On the score sheet a *one* or a *zero* was recorded under each category heading, for each problem, to indicate whether or not the behavior did occur. A total of 132 problems was scored for each of the 68 categories.

The in-baskets were scored by eight scorers, each one scoring a different half in-basket. Reliability was determined by correlating the scores obtained from the odd-item scorers with the scores for the even-item scorers, and correcting for length. Thus the reliabilities reflect both scoring accuracy and amount of consistency in the principals' behavior. The reliabilities which resulted varied from zero to .97, as shown in Table 1.

The 40 categories which are double-starred in Table 1 were chosen for use in the next phase of the analysis. These categories had reliabilities of .52 or higher and a median reliability of .78. The intercorrelations of these 40 scores were computed, and the matrix of intercorrelations was factored. Eight factors were retained; they account for almost all of the common variance among the scores. Rotations were made graphically to produce an oblique factor matrix with simple structure. Table 2 of the handout presents the category scores having loadings of .25 or more on each of the factors.

Factor A is called *Exchanging Information*. This interpretation seems quite clear on the basis of the four scores with the highest loadings.

Factor B we call *Discussing Before Acting*. Again the interpretation is clear from the scores with the highest loadings. The "new structure" in *initiates a new structure* is likely to be a committee or other discussion group; the "procedure" in *arrives at a procedure for deciding* is likely to be a discussion.

Table 1

Reliability of In-Basket Category Scores

Category	Reliability
**Estimated Number of Words Written	.94
**Number of Items Not Attempted	.97
**Number of Courses of Action (Usual)	.92
Rejection of Test Conditions	.59
**Number of Subordinates Involved as Individuals	.84
**Number of Subordinate Groups Involved	.69
**Number of Superiors Involved	.60
**Number of Outsiders Involved as Individuals	.73
Number of Outside Groups Involved	.44
**Unusual Action	.52
Recognizes Good Work	.71
**Aware of Poor Work	.65
**Carelessness or Minor Error	.69
Socially Insensitive	.38
**Relates to Background Material or Other Items	.75
**Conceptual Analysis	.70
Prejudges, Unwarranted Assumptions, Inappropriate	.35
Uses Human Values	.43
Uses Physical Values	*
**Uses Program Values	.65
**Discusses with Subordinates	.84
Discusses with Other Principals	.39
**Discusses with Superiors or Outsiders	.58
**Asks Subordinates for Information or Opinion	.81
Asks Superiors for Information or Opinion	.41
Asks Outsiders for Information or Opinion	.46
**Requires Further Information for Deciding	.68
**Delays, Postpones, or Temporizes	.81
**Arrives at a Procedure for Deciding	.80
Contingent Decision	.38
**Concluding Decision	.78
**Tentative or Definite Plans Only	.92
**Work Scheduled for Same or Next Day	.77
**Work Scheduled for Same or Next Week	.83
**Work Scheduled—No Time Specified	.83
**Leading Action	.90
**Terminal Action	.86
**Follows Lead by Subordinates	.77
**Follows Lead by Superiors	.71
**Follows Lead by Outsiders	.61
**Follows Pre-Established Structure	.67
Coordination	.41
**Initiates a New Structure	.78
Delegates Completely	.31
Delegates Partially with Control	.24
Delegates Partially without Control	.59
**Gives Directions or Suggestions	.83
Refers to Superiors	.17
**Communicates Face-to-Face	.86
**Communicates by Telephone	.74
**Communicates by Writing	.90
**Gives Information to Subordinates	.66
Gives Information to Superiors	.24
**Gives Information to Outsiders	.58
Explains Actions to Subordinates	.53
Explains Actions to Superiors	.19
Explains Actions to Outsiders	.18
**Courtesy to Subordinates	.91
Courtesy to Superiors	.53
**Courtesy to Outsiders	.55
**Informality to Subordinates	.92
Informality to Superiors	.64
Informality to Outsiders	.00
Backs up Staff	.06
Improves Staff	*
Improves Working Conditions	*
Sets a Deadline	.36
Follow-up or Feedback Planned	.01

*Reliability estimates were not computed because of the extremely low frequency of these scores.

**These items were used in the first factor analysis.

Factor C might appear to involve precipitate and aggressive action from the categories *concluding decision* and *terminal action*. But the two *following leads* categories also have high loadings. The actions and decisions are thus likely to be made in compliance with suggestions; hence the name *Complying with Suggestions Made by Others*.

Factor D has very high loadings on two categories. The factor is called *Analyzing the Situation* because it appears to involve broad situational analysis of in-basket problems.

Factor E has loadings on categories which imply concern about superiors and outsiders, and the negative loading on *delays* implies prompt action. We call this factor *Maintaining Organizational Relationships*.

Factor F is *Organizing Work*, and it is characterized particularly by care in specifying in advance quite exactly when one's work is to be done.

Factor G is called *Responding to Outsiders* because the four categories with the highest loadings all have to do with people outside the organization. We might think of this factor as reflecting concern about community relations.

Factor H is interpreted as *Directing the Work of Others*. Giving directions to subordinates is likely to be done in writing rather than orally, and courtesy is often used, apparently to soften the blow.

A second-order factor analysis revealed two factors which we call X and Y. The saturations of the individual scoring categories were determined, and those saturations of category scores which are .50 or higher are shown in Table 3.

Factor X has high saturations on a number of scores having to do with productivity—number of words written, number of courses of action taken, number of people involved, and so on. There are negative loadings on such categories as *delays or postpones* and *plans only*. We call this second-order factor *Amount of Work Done in Handling Items*.

Factor Y is a bipolar factor with negative loadings on *concluding decisions* and *terminal action* and positive loadings on a variety of scores having to do with deciding how to proceed, getting informed, and having discussions. Hence the name, *Preparation for Decision vs. Taking Final Action*.

The Factor Y loadings shown in Table 3 represent the ex-

Table 2

Factor Loadings

Factor A: Exchanging Information	
Asks subordinates for information, opinion or advice	.50
Gives information to subordinates	.45
Requires further information for deciding	.34
Gives information to outsiders	.31
Number of subordinate groups involved	.29
Number of usual courses of action	.25
Factor B: Discussing Before Acting	
Work scheduled—time unspecified	.63
Discusses with subordinates	.62
Communicates face-to-face	.60
Initiates a new structure	.48
Arrives at a procedure for deciding	.46
Tentative or definite plans only	.40
Number of items attempted	.40
Discusses with superiors or outsiders	.39
Number of usual courses of action	.36
Requires further information for deciding	.31
Number of subordinate groups involved	.29
Follows lead by subordinates	.29
Follows lead by superiors	.29
Terminal action	–.25
Factor C: Complying with Suggestions Made by Others	
Concluding decision	.73
Number of items attempted	.68
Follows lead by subordinates	.65
Terminal action	.59
Follows lead by superiors	.50
Follows pre-established procedure	.48
Communicates by writing	.45
Number of words written	.33
Number of subordinates involved	.31
Gives directions or suggestions	.31
Gives information to subordinates	.28
Number of usual courses of action	.27
Informality to subordinates	.27
Number of subordinate groups involved	.26
Factor D: Analyzing the Situation	
Uses program values	.82
Conceptual analysis	.75
Aware of poor work	.28
Factor E: Maintaining Organizational Relationships	
Number of superiors involved	.54
Discusses with superiors or outsiders	.42
Number of outsiders involved	.41
Relates to background materials or other items	.37
Follows lead by outsiders	.26
Communicates by telephone	.26
Delays, postpones	–.38
Factor F: Organizing Work	
Work scheduled for same or next week	.51
Work scheduled for same or next day	.50
Follows pre-established procedure	.30
Relates to background information or other items	.25
Work scheduled—time unspecified	–.48
Factor G: Responding to Outsiders	
Gives information to outsiders	.44
Courtesy to outsiders	.44
Follows lead by outsiders	.41
Number of outsiders involved	.31
Carelessness or minor error	.30
Awareness of poor work	–.31
Factor H: Directing the Work of Others	
Leading action	.64
Communicates by writing	.59
Courtesy to subordinates	.57
Gives directions or suggestions	.47
Courtesy to outsiders	.40
Carelessness or minor error	.31
Number of subordinates involved	.28
Communicates by telephone	–.30
Tentative or definite plans only	–.67

tremes of a continuum which has a rather marked similarity to certain theories of decision-making. Table 4 identifies six theoretical stages in the process of decision-making. In the parallel columns are shown some selected category scores and their loadings on Factor Y. The agreement between the theoretical formulation and the empirical findings is rather striking. This is, of course, no verification of the theory that one goes through these stages in reaching a decision. There is clear indication, however, that a principal's characteristic behavior in response to a standard set of administrative problems can be described in terms of his position on the continuum of decision development.

A large number of other variables—ratings, inventory scores, ability measures, and so on—was also available. The relation of the factor scores to these other variables is a matter of considerable interest. Therefore the intercorrelations of 120 variables were computed. The variables included in-basket category scores, ability measures, personality inventory scores, ratings, interest measures, tests of professional knowledge, and biographical items.

It is, of course, possible merely to compute the correlations between factor scores and other variables; but such an approach is possibly misleading because of the fact that the

Table 3

Second-Order Factors

Factor X: Amount of Work Done in Handling Items	
Number of words written	.67
Number of usual courses of action	.65
Number of outsiders involved	.60
Gives directions or suggestions	.57
Number of subordinates involved	.57
Communicates by writing	.56
Leading action	.53
Gives information to subordinates	.52
Follows lead by superiors	.51
Factor Y: Preparation for Decision vs. Taking Final Action	
Arrives at a procedure for deciding	.69
Requires further information for deciding	.64
Work scheduled for same or next day	.61
Discusses with subordinates	.60
Asks subordinates for information or advice	.56
Communicates face-to-face	.55
Initiates a new structure	.54
Work scheduled for same or next week	.54
Concluding decisions	−.52
Terminal action	−.62

factors are substantially correlated and hence contain variance which is shared by other factors. A factor-analytic approach to the problem was adopted, at the suggestion of Dr. Ledyard Tucker, which makes it possible, in effect, to determine the relationship of each variable to the part of each factor which is not shared by other factors. For example, *Amount of Work* influences in varying degrees the scores on all the eight factors; it would be desirable to learn the relationship between variables and factors with the effect of *Amount of Work* ruled out. The method employed resulted in the computation of coefficients which are proportional to the correlations with the unique part of each of the eight factors.

The procedure is as follows: the 120 x 120 matrix was factored. An orthogonal factor matrix composed of the first ten factors was rotated to form an oblique matrix having a

Table 4

Stages in Decision-Making

Stage in Decision-Making	In-Basket Score	Loading on Factor Y
1. Recognizes a problem		
2. Prepares to clarify the problem	Arrives at a procedure for deciding	.69
	Requires further information for deciding	.63
3. Initiates work on the problem	Work scheduled for same or next day	.61
	Discusses with subordinates	.60
	Asks subordinates for information	.56
	Leading action	.41
4. Organizes and judges facts and opinions	Conceptual analysis	.24
	Tentative or definite plans only	.16
	Delays, postpones	.15
5. Selects alternative solutions	Follows lead by superiors	.09
	Follows lead by subordinates	−.10
6. Decides and acts	Concluding decision	−.51
	Terminal action	−.62

Table 5

Relationships of Cognitive Ability Tests to In-Basket Test Factors

	1 Exchanging Info.	2 Discussing	3 Complying w/ Sugg.	4 Analyzing Situation	5 Maintaining Rela.	6 Organizing Work	7 Resp. to Outsiders	8 Directing	9 Amount of Work	10 Prep. for Decision
Reasoning	.02	.03	.48	.34	−.15	.07	−.31	−.22	.28	.13
Subtraction and Multiplication	.43	−.14	.00	−.25	.14	.09	.12	−.26	.38	.21
Vocabulary	.39	−.14	.06	.15	.03	−.06	−.12	−.31	.33	.24
Concealed Figures	.04	.16	.47	.22	−.22	.11	−.13	−.24	.30	.24
Mathematics Aptitude	.06	.04	.49	.24	−.13	.18	−.17	−.30	.34	.16
School Administration and Supervision	.36	−.09	.15	.18	.01	.04	−.23	−.28	.41	.45
NTE Science and Mathematics	.07	.09	.48	.33	−.21	.02	−.27	−.27	.28	.19

factor structure as nearly as possible like that found for the original in-basket factor analysis. Coefficients were computed which reflect the relative relationship of each of the 120 variables to each of eight oblique reference vectors, each vector corresponding to one of the eight in-basket factors. These coefficients are proportional to the correlations with the unique portion of the corresponding factor. Similar estimates of correlations with second-order factors were computed. We will have time merely to indicate briefly a few salient findings.

Table 5 shows the relationships of several cognitive measures with in-basket factors. We see that both of the second-order factors (Columns 9 and 10) have fairly high relationships with the tests. *Amount of Work* is related to a variety of cognitive abilities, while *Preparation for Decision* relates especially to the School Administration test. The signs of the correlations in Column 10 show that high ability people tend to prepare for decision rather than to take action.

The other columns show relationships with the unique parts of the eight in-basket factors—the part remaining after the variance due to second-order factors is removed. The coefficients in the first eight columns therefore cannot be attributed to *Amount of Work* or to *Preparation for Decision.*

There are a number of high positive coefficients, particularly in Columns 1 and 3. People who characteristically exchange information (Column 1) are high on Verbal and Number factors and on the test of school administration. The relationships with Complying with Suggestions (Column 3) involve a completely different set of tests, particularly tests of reasoning. Perhaps the compliance factor is not weak submission, but involves logical evaluation of suggestions.

Coefficients in Columns 7 and 8 are mostly negative. Those principals who were responsive to outside pressures and who were characterized by actively directing the work of their subordinates tended to be the less able principals.

Table 6 gives relationships between in-basket factors and some selected scores from Cattell's 16 Personality Factor Inventory. The personality scores do not predict the second-order factors; but there are a number of high relationships with the unique parts of the primary factors. Look at Column 5, for example. Principals who typically try to maintain good

Table 6

Relationships of Personality Measures to In-Basket Factors

	1 Exchanging Info.	2 Discussing	3 Complying w/ Sugg.	4 Analyzing Situation	5 Maintaining Rels.	6 Organizing Work	7 Resp. to Outsiders	8 Directing	9 Amount of Work	10 Prep. for Decision
Friendly vs. Aloof	.21	−.09	−.29	−.28	.51	−.14	−.15	−.18	−.01	.01
Emotional Maturity vs. Lack of Frustration Tolerance	.15	.28	−.09	−.10	.12	−.36	.08	−.03	.12	.07
Dominance vs. Submission	−.14	.03	.12	.30	.14	.03	−.37	−.12	−.03	.09
Adventurous vs. Shy	.03	.19	−.26	.01	.50	−.33	−.23	−.07	−.02	.08
Emotional Sensitivity vs. Tough, Practical	.34	−.07	−.21	−.28	.22	−.11	.12	−.19	.10	.05
Suspicious vs. Trusting	−.30	−.16	.36	.18	.14	.45	−.13	.00	−.02	−.13
Sophistication vs. Rough Simplicity	−.12	−.10	.09	.30	.02	−.02	−.35	.01	−.08	−.03
Anxious Insecurity vs. Placid Self-Confidence	−.21	−.28	.25	.14	−.35	.55	−.01	.03	−.11	.00
Nervous Tension	−.21	−.30	.39	.22	−.30	.60	−.22	−.14	−.10	.04

Table 7

Relationships of Biographical Information to In-Basket Test Factors

	1 Exchanging Info.	2 Discussing	3 Complying w/Sugg.	4 Analyzing Situation	5 Maintaining Rels.	6 Organizing Work	7 Resp. to Outsiders	8 Directing	9 Amount of Work	10 Prep. for Decision
Experience in Education	.09	−.18	−.49	.01	.08	.04	.44	.28	.04	.01
Years of Academic Preparation	−.01	.02	.08	.08	−.08	−.01	−.15	.04	.03	.14
Age	.02	−.15	−.43	.05	.06	.02	.40	.28	.00	−.05
Sex (Male = 1, Female = 2)	.45	.17	−.60	−.42	.22	−.07	.50	−.03	.08	.13

organizational relationships tend to be friendly, adventurous rather than shy, and free from anxiety and nervous tension. Look at Column 6. Principals who typically plan their work for specific days and hours tend to lack frustration tolerance and to be shy, suspicious, anxious, and nervous. Anxious people appear to exhibit a compulsive pattern of behavior in handling in-basket problems.

Table 7 shows relationships of in-basket factors with some biographical information items. Again the second-order factors are unrelated to the predictors, but the unique parts of the primary factors have high relationships with certain of the items. Years of college training has nothing to do with any of the factors; but age, experience, and sex do. *Complying with Suggestions* (Column 3) is typical of inexperienced young men. *Responding to Outsiders* (Column 7) is characteristic of experienced old women.

The simulation of a standard job in educational administration through the use of in-baskets has proved to be successful as a method of collecting records of administrative performance which can be scored reliably, and which yields scores which are useful in providing a better understanding of some of the dimensions of performance in such a situation. The method of factoring which was employed appears to be a powerful technique for isolating important aspects of behavior and examining their relationships with other measures.

ROBERT M. W. TRAVERS, *Department of Educational Psychology, University of Utah*

Models of Teacher Behavior in the Classroom

THE BUILDING of models, supposedly descriptive of what happens in American classrooms, has become an activity which occupies the time of numerous persons from many different sections of our culture. Most educational reformers begin with the premise that certain activities, of which they disapprove, occupy much of the time of pupils in the classroom, and that changes should be made which would eliminate these activities. Those who believe that the schools are dominated by what they call a life adjustment curriculum fall into this category of imaginative model builders. In addition, there are theorists among the behavioral scientists who also have models based on certain beliefs concerning what goes on in classrooms. They also advocate certain reforms designed to change what they imagine to occur in classrooms. Most of these model builders base their descriptions of what supposedly occurs in classrooms on hearsay, an examination of textbooks read by students of education, and anecdotal material brought home by their children. A few of the more eager ones make a few visits to schools. Enthusiasm for describing and attacking what goes on in hypothetical schools has hardly been matched by enthusiasm for collecting data concerning what actually happens. Surely educational reform should begin, not with models of classroom happenings based on imagination, but on carefully and perhaps tediously collected data concerning what actually happens in American classrooms. My purpose here is to present a small amount of data concerning this matter based on what must be judged to be a limited amount of observation—namely about 500 hours.

The data which are reported here are a by-product of a research project which my colleague, Dr. Norman Wallen, and

I have been conducting as a part of the Cooperative Research Program of the United States Office of Education. Our project is concerned with the determinants of teacher behavior and the data collection process has taken us into elementary schools where we have spent many hundred hours during the last year recording the events that take place. Our project is not concerned in any way with the evaluation of teachers but is designed to study teaching as a natural phenomenon so that some of the variables related to particular patterns of teacher behavior may be identified.

An inevitable result of this study is that a considerable volume of data has been collected concerning the practices and procedures of teachers in the schools in which the research has been conducted. While the primary interest in such data is a scientific one, the material is also of general interest at this time since the public has become deeply concerned with what is happening in schools. The data presented here represent a small contribution of knowledge concerning actual events in schools which may be of interest to reformers and a starting point from which plans for educational change might begin. While our data are limited, they are still extensive compared with most other data concerning what actually happens in schools.

Our data were collected in five schools within the Salt Lake City School District. These schools are located in a variety of different social-economic neighborhoods as well as in different parts of the town. The schools were not originally selected to be representative of the school system although they are to a considerable degree. The only known differences between these schools and the other schools of the system rest in the fact that they represent some of the larger units. In addition, for the purposes of our research, we wished to include some teachers who were known to represent the progressive tradition and for this reason we included the five lower grades in the campus school at the University of Utah. The inclusion of the latter grades probably loads our sample rather heavily with teachers who state that they represent a progressive tradition. This is an important point to keep in mind.

The data to be considered here consist of samples of statements made by 83 teachers during two separate visits conducted by each of two observers, Dr. Wallen and myself. Our procedure was essentially that developed by Withall (1) which required the observer to record the first statement made during each minute interval until twenty-five statements had been collected. The only types of verbal behavior excluded from our observation were certain repetitive verbal behaviors as would occur, for example, when the teacher was reading a spelling test. Although an attempt was made to collect 25 statements during each visit to a teacher, this was not always feasible. Some teachers showed relatively little verbal behavior and several hours of observation were necessary in order to obtain two samples of twenty-five statements. Since the two observers each collected two samples from each one of the teachers, a total of 100 samples of verbal behavior were collected on each teacher. Altogether, on the 83 teachers, a total of 8,300 samples of verbal behavior was collected.

Precautions were taken to ensure that the introduction of an observer into the classroom would have as little influence as possible on the teaching process. The teachers were given a broad orientation concerning the purposes of the research and were assured that they were not being rated for competence. They were also assured that all records would be considered confidential and that the results of the study would be reported back to the participating teachers. They were also asked to conduct their classes as planned when the visitor arrived and not to put on any special program or demonstration at the time of the visit. Preliminary visits were made to each classroom until the teacher seemed at ease with an observer. An effort was made to know the teachers on a personal basis by chatting with them during the recess and before class in the morning. A great amount of time and effort was devoted to developing good relations with teachers so that they could be at ease during the observation sessions. The project staff regarded this as an excellent investment of time.

Our general observation supports the point of view that these samples of verbal behavior are to a high degree representative of the behaviors manifested by the teachers. Some aspects of behavior are inevitably missed by the technique. A glance of approval or disapproval would not be recorded nor would a teacher-to-pupil remark made in a low whisper. Our impression is that most of the behavior of most teachers is public verbal behavior.

One does not have to make an analysis of all the 8,300 items of verbal behavior to obtain an over-all impression of what happens in these classrooms. For this purpose a sample will suffice. Hence we have selected from each group of 25 verbal behaviors a sample of one which in each case was selected through the use of a table of random numbers. This provides a group of 332 verbal statements distributed over 83 teachers. A classification of these statements is given in Table 1.

Table 1

Distribution of Sample of 332 Verbal Behaviors of Teachers

Category of Behavior	Frequency	Percent of Total
Informing (academic)		
Emits information	30	9.0
Demonstrates		
Provides cues	1	0.3
Corrects	7	2.1
Asks questions	50	15.1
Reinforces	16	4.8
Relates personal experience	1	0.3
Gives doubtful reinforcements such as "Yes"	2	0.6
Controlling		
Reprimands	19	5.7
Expresses frustration		
Exercises control to gain recognition		
Asks class or pupil to be quiet	7	2.1
Setting of goals and standards and initiating activity		
Asks pupil or pupils to exercise choice	5	1.5
Directs pupils what to do or what goal to achieve	72	21.7
Gives or scores tests or checks up on pupil	9	2.7
Management		
Performs routine class management functions	41	12.3
Delegates management function to pupil	3	.9
Organizes games, recreations, rest periods	5	1.5
Reads (literary)	1	0.3
Seeks administrative information	16	4.8
Emits administrative information	24	7.2
Affiliation		
Expresses affection		
Expresses need for good social relations		
Rejects		
Suggests one child help another	1	0.3
Small talk	7	2.1
Unclassifiable		
Ambiguous	15	4.5

There are a number of matters of interest to point out in this table. First, consider the total frequency of occurrence of verbal behaviors in each one of the categories. The most frequently occurring is that of telling the pupil what to do, a traditional practice which has had many advocates through the ages. Of course, many contemporaries who think about educational problems believe that more of pupil behavior should be self initiated, but the fact is that in our sample the teachers devoted much of their energy to directing the academic work of the pupils. A second point to note is that the next most frequently occurring form of teacher behavior is that involved in a questioning process. This reflects the extensive use of the traditional recitation procedure. The third most frequently occurring form of teacher behavior, outside of performing classroom management functions, is that of providing information. This also is a traditional function of teachers, although one about which many questions have been raised in recent educational literature. In contrast, the categories of teacher behavior which the progressive education movement of the thirties considered to be important show low or zero frequencies of occurrence. For example, it is a rare event for a teacher to delegate to a pupil a decision-making function. Only 1.5 percent of the teacher behaviors fall into this category. The emphasis on personal and social adjustment, which some claim to be a focal weakness in our educational system, is completely lacking in our data. Indeed, the impression given by the table is that of a subject matter-oriented curriculum taught by methods which have been traditionally practiced and which the public widely accepts.

The data agree well with our general observations. We noted a high degree of emphasis on the learning of subject matter. The capacity of teachers to handle a range of subject matter which covered such diverse topics as meteorology, physics, physiology, geology, as well as the more conventionally taught areas of knowledge was impressive. The excellent preparation of the teachers in this respect is demonstrated by the fact that out of 8,300 recorded statements of teachers, there is only one subject matter error. This is an enviable record and one which presents a standard which I would like to live up to in my own teaching though I am sure I do not. The traditional subject matter of education is amply represented in our sample. If two expressions giving an over-all picture of the classrooms of these teachers were to be chosen, I think the two most appropriate would be that they were subject matter-oriented and businesslike.

Psychologists, who have recently joined the ranks of educational reformers, are likely to look at the data much as they have sometimes looked at classrooms and jump to the conclusion that teaching does not provide adequate reinforcements for efficient learning. I do not think that such a conclusion is justified. The reinforcements operating in the classroom are very difficult to identify. Indeed, until I had spent a few hundred hours in classrooms, I did not become aware of the wealth of reinforcements that are operating. For example, in most classrooms if a teacher does not comment on a student's answer, the pupil will assume that his answer was right. Silence is probably the most commonly occurring reinforcer and hence does not appear in our data. Other reinforcers, which are not easily observed, are those built into the books and workbooks. A child stumbling through an elementary reader is reinforced for those efforts which result in the emergence of a meaningful story. Again, many cues may be present which indicate to a child solving a mathematical problem that he is right. Psychologists who visit classrooms and report an impoverishment of reinforcing agents have been looking in the wrong direction. I suspect that psychologists are prone to look for the kinds of reinforcing agencies that they commonly introduced into the laboratory and since they cannot see these in the classroom, may assume that few reinforcements are provided. Having arrived at this conclusion the next step is likely to be that of designing a classroom which resembles their own laboratory and in which they feel comfortable. This kind of educational reform is analogous to the efforts of an imaginary psychologist who, after spending a lifetime studying the behavior of chickens, spent his years of retirement writing a book on how to teach rats to peck.

While the data refer to a single school system, my impression is that these classrooms are very similar to those which I have visited elsewhere. This suggests to me that teaching is performed largely in terms of a cultural tradition. Teachers have to satisfy the expectations of parents, children, school boards, and other community groups, if they are to retain their jobs, and these expectations may well play a much more significant role as determinants of teacher behavior than do all of the verbal transactions involved in teacher education. Our data are not consistent with the notion found in popular magazines that the ideas of John Dewey have produced a radical change in teaching practices. As a matter of fact, the influence of John Dewey is difficult to discern in our data.

There is no evidence in our data to support the notion that the teachers in our sample were particularly concerned with a life adjustment curriculum of the type that has been criticized in the press. The behavior of our teachers shows a central concern for intellectual development, and one has to look hard among our recorded 8,300 verbal statements to find some which have to do with the emotional and social development of pupils. Certainly less than one-half percent of teacher behavior recorded in our study has anything to do with this objective. The rare teacher behaviors that have to do with this objective were generally initiated by some crisis such as occurred in one classroom where the teacher noted that a foreign child who did not speak English was being mistreated by his classmates. The educational process, as we have observed it in our sample of teachers, was almost exclusively intellectualistic in both purpose and execution, and this despite the fact that we went to considerable pains to locate and include some teachers who supposedly taught in a so-called progressive tradition.

Our data leave us with the impression that the behavior patterns of the teachers in our sample represent a teaching tradition which is probably highly stable and difficult to change. The influence of teacher education is seen largely in command of subject matter, an area in which our teachers showed extraordinary competence. Those aspects of teacher education which have to do with teaching methods show much less discernible effects, which perhaps is hardly surprising in view of the fact that such courses generally represent only about 25 percent of the teacher-training curriculum. Also, professional courses in education are faced with the very difficult task of changing already well-established concepts of education and habits of dealing with the young. I am not saying that professional courses in education are not important, for I am convinced that they are, but they attempt to undertake a task which is fantastically difficult to accomplish.

The data which we have presented and the total body of data from which it is derived, represent only a small portion of the data needed for the building of what one might hope to be a useful model of the behavior of teachers as it occurs in schools. There is a great lack of data concerning the incidence and operation of reinforcements in the classroom and the extent to which reinforcing functions are and can be delegated to workbooks and other forms of equipment. Data on this problem may well have to be collected in miniature experimental situations rather than in the classroom itself. We look forward to having available an increasing body of data which can be used to develop a theoretical model of how teachers operate. Such knowledge, as it evolves, will slowly equip the scientist with some of the knowledge which he must have if he is to introduce successful changes into the system.

REFERENCE

1. Withall, John. The development of a technique for the measurement of social emotional climate in the classroom. *Journal of Experimental Education*, 1949, 16, 347-361.

CARSON McGUIRE,
Department of Educational Psychology, The University of Texas

The Prediction of Talented Behavior in the Junior High School

A TALENT is a pre-eminent aptitude, or a superior ability, either natural or acquired, or a capacity for achievement or success. In school settings, "human talent" refers not only to the various kinds of scholastic aptitude and academic achievement measured by objective instruments and evaluated by teachers, but also to other forms of behavior appraised by one's age-mates and older people. Thus, talented behavior involves both personally-significant and socially-valued competencies, including signs of creativity, recognized as such through performances or products which can be assessed by other persons.

The basic assumption underlying the Human Talent Research Project is that, if factors such as innate ability be held constant, talented behavior is acquired and becomes organized, or structured, and to some extent predictable, as a consequence of the educative process. Unless something out of the ordinary is done in guiding learning experiences, a process we term "telesis," the behavior of most individuals attains a kind of crude stability or invariance which makes it predictable. This paper presents an approach to studying and testing ideas about talented behavior in school and, eventually, in experimental settings. The major outcome of the research has been the development of "factor variables" which

---

Note: The research reported herein was supported through the Cooperative Research Program of the United States Office of Education, Department of Health, Education and Welfare. In this first stage of the Human Talent Research Project, faculty associates were Ralph L. Duke, Benjamin Fruchter, Beeman N. Phillips, John Pierce-Jones, and Jackson B. Reid; research associates on the project staff were Edwin Hindsman, Earl Jennings, F. J. King, and George McBee.

combine indicators of talented behavior. Two sets of factor variables, one for the general population with modifiers for the two sexes and the other found among highly-intelligent students, have turned out to be surprisingly stable and valid predictors with a good deal of explanatory value.

Nature of Talented Behavior

A general proposition about the nature of talented behavior provides a research model set forth schematically in figure 1. Talented behavior may be regarded as a function of:

(a) pertinent cognitive, perceptual, and psychomotor abilities (Bloom, 1956; Ferguson, 1956; Guilford, 1956, 1959; Wilson, 1958);

(b) elements of personality and motivation, especially expectations about one's behavior and the probable responses of other persons to it (Atkinson, 1957; Sears, 1951; Stern, Stein, & Bloom, 1956);

(c) ways in which an individual is valued and has pressures imposed upon him or her by others, such as parents, age-mates, or teachers, in various settings (Elkins, 1958; Haggard, 1957; McGuire, 1956).

These three classes of variables may combine in different ways as a consequence of (d) sex-role identification and the sex-typing of socialization pressures, and (e) the context of behavior, especially the varying patterns of educational experiences and role expectations in different school locations (Goethals, 1958; Ferguson, 1954, More, 1955).

Gathering and Processing the Data

Frankly speaking, the research team faced a monumental task of gathering, organizing, processing, and making sense of a wealth of data in three years, from the early fall of 1957 to the late spring of 1960. Decisions as to what instruments were to be administered were made by a committee in the spring and summer of 1957. A number of tests were revised for use with adolescents and other devices were prepared (McGuire & Associates, 1960). The administration of instruments was carried out largely by members of the project staff and faculty associates with some valuable cooperation from the people in the four school systems.

By means of several kinds of multivariate analysis in pilot studies, variables judged to be most pertinent to the study of talented behavior were identified (Hindsman & Duke, 1960). The team used IBM cards and, later, tapes to organize and store data. Programs were written for basic processes such as transforming the original values punched on cards after scoring into another deck of cards with stanine values. The reduction of raw data involving 140 measures into stanine values facilitated variance, factor-analytic, and multiple-regression analyses using IBM 650 equipment.

Table 1

Regression of GPA Teacher Evaluations for Grade VII upon Selected Indicator Variables by Community and Sex Role

(Decimal points omitted)

Indicator Variables[a]	Communities A M	A F	B M	B F	C M	C F	D M	D F
Sample population, N	161	157	94	112	266	241	112	113
Regression variance, R^2	.72	.59	.72	.70	.64	.68	.65	.71
Multiple correlation, R	.85	.77	.85	.84	.80	.82	.81	.84
1. CTMM Mental Function			19	10	18	13	35	10
3. DAT Mech. Reasoning			-06		-10		04	
4. Gestalt Transform'n	04		14	08	03	13	17	07
5. Rhymes (word fluency)	26	08	06	06	11	10	03	23
6. Unusual Uses			05	17		09	03	-25
7. Consequences			-15	-06	-03	05	-13	-10
8. Common Situations		-11	-03	-04	17		06	06
9. Seeing Problems	08	13	17	16		15	11	18
10. Mutilated Words	-03	08	-06	15	10	09	04	12
11. Gestalt Completion	-11	-09	10	12	-06	03		
12. Short Words	11	17	17	07	06	10	-06	-16
13. Dotting	-10	05	-10		07	-03	04	16
14. DRT Reaction Time	09			-04	07	-05		16
15. JPQ 1 Sensitivity			02	03	-03		10	-08
16. JPQ 3 Emotionality	-11		-10			-03	-10	06
17. JPQ 6 Sociability	04			03	-06	07	-08	-06
18. JPQ 8 Soc. Morale	-05		-08			-02	14	-11
19. JPQ 9 Dominance	10	-06	06	-06		03	15	-16
20. JPQ 11 Surgency		-10	07	-15		03	05	10
21. SSHA Schol. Motiv'n	10	23		12	20	12	14	26
22. CMAS Anxiety	-06		-19		04			-10
23. CYS Authoritarian			11	-20	-04		03	-13
24. CYS Critic. of Youth	06	-09	13	11			-04	
25. CYS Critic. Education	-11		-07	03	02		07	-15
26. CYS Neg. Orientation		07			-07		-03	
27. CYS Maladjustment		-09					-11	18
30. ISS Family Status			12	-06	-04	-04	04	18
31. N-1 Peer Acceptance	29	23	06	09	28	23	24	39
32. N-2 Model Value	17	07	-08	07	19	07		16
33. N-3 Effectiveness	19	20	18	19	20	14	13	-04
34. N-4 Nondeviant	05	15	20	07	08	05	12	
35. N-5 Quiet Dependency	09	-02	04		14	05	10	

[a] Beta weights with zero values have been left blank.

Then the most stimulating and productive steps in the processing of data began. By means of a series of factor-analytic studies of measures from 1,417 boys and girls in the seventh grade, we were able to map out the different ways in which variables combined to sort out and portray achievements, abilities, and attributes. We also sorted out the meanings of descriptions age-mates gave of one another in response to 46 nomination items. The analyses indicated that five sociometric factors were common to both sexes and, consequently, we could use a smaller number of factor scores to represent appraisals by peers. Repeated measurements in grades eight and nine for selected variables provided distributions for

Table 2

Multiple Correlation Coefficients for Criterion Measures

(Decimal points omitted)

Regression Analysis[a]		Criterion Measures					
		36 GPA Year	37 CAT Read	38 CAT Lang	39 CAT Arith	40 STEP So. St.	41 STEP Science
Community A	M	85	82	81	78	80	79
	F	77	84	82	77	84	77
Community B	M	85	83	82	82	84	85
	F	84	84	84	77	82	78
Community C	M	80	81	81	78	79	79
	F	82	85	83	81	86	85
Community D	M	81	81	75	79	79	77
	F	84	83	81	78	85	81
Males (N = 634)	(1)	75	78	75	71	76	75
	(2)	75	78	73	70	80	79
	(3)	74	71**	68	66	71**	71**
Females (N = 608)	(1)	75	81	77	72	81	77
	(2)	75	81	77	72	82	82
	(3)	74	75**	73	67	76**	75**
Total Population (N = 1242)	(1)	74	78	75	69	77	73
	(2)	75	78	75	70	80	80
	(3)	74	72**	71	63**	72**	72**

[a] Each coefficient represents a separate regression analysis illustrated in Table 1. The three types of combined analyses are:

(1) CTMM Mental Function included, STEP Listening omitted;
(2) STEP Listening included, CTMM Mental Function omitted;
(3) Multi-factor cognitive variables omitted; namely, CTMM Mental Function, STEP Listening, DAT Mechanical Reasoning, Rhymes.

** Multiple correlation coefficient significantly different at the .01 level of confidence from the largest coefficient for the same criterion measure and population with a different set of predictors (F test).

studying reliabilities as well as criterion measures for the prediction studies (McGuire & Associates, 1960).

Two complementary techniques, multiple regression and factor analysis, were employed to identify the significant indicators of talented behavior and to reduce them to an even smaller number of meaningful factor variables. As shown in Tables 1 and 2, multiple correlation coefficients ranging from $R = .75$ to $R = .86$ were obtained to represent the relationships between criterion measures, such as teacher evaluations and standard tests, to various combinations of the 32 indicator variables. Table 2 would lead one to infer that measures of strength of association were not quite so high when data were combined across communities for boys and girls. Nevertheless, the analyses almost invariably explained 50 per cent or more of the variation in the different kinds of valued performances.

The results encouraged the research team to take the next step of reducing the 32 promising indicators to a smaller number of factor variables which could more effectively explain, yet still predict with some precision and confidence, different kinds of talented behavior. As summarized in Table 3, this was done first for the total population of 1,242 students who had completed the necessary instruments in grade seven. In the general population, we sorted out five factors common to boys and girls, two specific to each sex, and two characterizing deviant behavior. Table 3 shows the factor loadings and the optimal weights for the component indicators of talented behavior, the latter being determined by a multiple regression method. Using the optimal weights and a suitable program, new sets of factor scores were computed for each individual to be considered in the prediction studies.

Development of Factor Variables

Techniques employed in the development of factor variables (Cattell, 1957, pp. 287-296; Guilford, 1954, pp. 524-526) can be illustrated using data gathered on 213 students of high mental function (HMF) in the four school locations. This study of highly-intelligent or potentially-gifted boys and girls, all one standard deviation above the mean on the CTMM, was designed to follow up an earlier inquiry into the behavior

Table 3

Optimal Weights of Indicators for Factor Variables

(Rounded to two decimal places)

Factor Variable	Indicator of Talented Behavior	Factor Loadings	Optimal Weights
	Common to both Sexes		
A. Cognitive Approach	STEP Listening	.73*	.26
(CA)	CTMM Mental Function	.69*	.21
	Gestalt Transformation	.60*	.19
	DAT Mechanical Reasoning	.59*	.19
	Rhymes	.59*	.09
R = .88	Gestalt Completion	.47*	.09
	ISS Family Status	.45*	.07
	DRT Reaction Time	.41*	.07
	Unusual Uses	.57*	.07
B. Divergent Thinking	Consequences	.61*	.32
(DT)	Seeing Problems	.61*	.32
R = .75	Common Situations	.60*	.28
C. Socially-Oriented	JPQ 8: Socialized Morale	.66*	.31
Achievement Motivation (AM)	SSHA Scholastic Motivation	.60*	.22
	JPQ 3: Emotional vs. Stable	−.60*	−.20
	JPQ 9: Independent	−.52*	−.17
R = .85	CYS Criticism of Education	−.51*	−.17
	CMAS Anxiety	−.52*	−.17
	JPQ 6: Sociability	.39*	.11
D. Peer Stimulus Value	Nom 1: Peer Acceptance	.72*	.50
(SV)	Nom 3: Effectiveness	.59*	.25
R = .78	Nom 5: Quiet Dependency	−.45*	−.23
E. Absence of Negative Valuations (NV)	Nom 2: Absence of Negative Model Value	.63*	.46
R = .72	Nom 4: Nondeviant vs. Deviant Behavior	.59*	.39

Table 3 (continued)

Specific to Boys			
F. Anxious Emotionality (AE)	CMAS Anxiety	.68*	.45
	CYS Personal Maladjustment	.57*	.29
R = .77	JPQ 1: Sensitivity	.33*	.20
	JPQ 3: Emotionality	.43*	.12
G. Anti-Social Wariness (AW)	CYS Authoritarian	.30*	.42
	CYS Neg. Soc. Orientation	.51*	.22
	CYS Critic of Education	.31	.21
R = .67	Short Words	.31*	.18
	Mutilated Words	.30	.16
Specific to Girls			
H. Sensitive Dependency (SD)	CYS Authoritarian	.43*	.34
	JPQ 1: Sensitivity	.45*	.33
R = .67	JPQ 8: Socialized Morale	.40	.26
	JPQ 11: Surgent/Desurgent	−.35*	−.15
I. Perceptual-Motor Skills (PM)	Gestalt Completion	.39	.25
	CYS Neg. Soc. Orientation	−.29	−.19
	Dotting	.31	.16
R = .48	DRT Reaction Time	.28	.11
Factor IV for Total Population			
J. Practical-Minded Toughness (PT)	JPQ 1: Sensitive vs. Tough	−.56*	−.42
	DAT Mechanical Reasoning	.36	.28
	Mutilated Words	−.34	−.26
R = .71	Short Words	−.31*	−.18
	JPQ 8: Socialized Morale	−.31	−.08
Factor VI for Total Population			
K. Anxious-Dependent Resentment (AR)	CMAS Anxiety	.46	.30
	CYS Authoritarian	.46*	.27
R = .68	CYS Neg. Soc. Orientation	.49*	.26
	JPQ 11: Surgent/Desurgent	−.29*	−.25

* Highest loading for the indicator variable in a factor structure.

and achievement of gifted children by Haggard (1957). The first step was to construct a master intercorrelation matrix for 32 promising indicators of talented behavior and six criterion measures, using values obtained in grade seven. As shown in Table 4, the regressions of criterion measures upon potential indicators were computed using an iterative technique (Greenberger & Ward, 1956) which had a stop criterion to avoid overfitting the regression line. Because original stanine values were employed, most of the cognitive distributions were restricted in range and the obtained multiple correlation coefficients were lower than for the general population.

Since the indicators were represented in sets of predictors for criterion measures, all 32 were retained for the development of factor variables. Correlations among the 32 indicators provided a matrix for the extraction of centroid factors which were rotated to an orthogonal normal varimax solution (Kaiser, 1958) by appropriate programming. The varimax criterion for analytic rotation was chosen since it had the greatest likelihood of portraying factors invariant with changing samples of tests and populations. As shown in Table 5, the analysis yielded seven factors which are described more fully in Table 6. The latter table was compiled to show the provisional names of the HMF factor variables, the significant indicators for each factor with their loadings, the optimal weights, and the multiple correlations between factors and sets of indicator variables. Optimal weights for the estimation of factors in persons were obtained by the multiple regression method so as to remove linear restraints (Guilford, 1954, p. 404).

The resulting factor variables were not entirely independent of one another, as shown in Table 7 where $r = .1765$ is significant at the .01 level of confidence. Nevertheless, none of the five significant correlations among the 21 represented more than ten per cent common variance between factors. Hence linear restraints were regarded as being minimal for multiple regression and prediction studies. Among the HMF factor scores, only the first had no significant association with a measure of academic performance according to the correlations summarized in Table 8.

Table 4

Regression of Criterion Measures of Talented Behavior on 32 Indicator Variables for 213 Junior High School Students of High Mental Function

(Rounded values, decimal points omitted)

Indicators	Teacher GPA	CAT Reading	CAT Language	CAT Arith.	STEP Soc. St.	STEP Science
1. CTMM Mental Function	19	47	39	38	10	14
2. STEP Listening	07	29	09	06	29	30
3. DAT Mechanical Reasoning	−04	−03	07	17	06	18
4. Gestalt Transformation	..	07	−04	..	15	..
5 Rhymes (verbal facility)	..	13	00	06	09	06
6. Unusual Uses	10	−08	03	−07	19	04
7. Consequences	11	..	10	09	..	..
8. Common Situations	−04	..	−27	−11	−10	..
9. Seeing Problems	..	07	03	..	09	−12
10. Mutilated Words	13	−07	16	07	−09	11
11. Gestalt Completion	..	..	−06	−09	..	−11
12. Short Words	06	15	19	22	08	..
13. Dotting	05	..	03	−08	−09	..
14. DRT Reaction Time	..	..	−04	05	09	08
15. JPQ 1: Sensitivity	07	..	..	03	04	−03
16. JPQ 3: Emotionality	05	03	10	08	−03	07
17. JPQ 6: Sociability	−03	..	04	12	−03	04

18. JPQ 8: Social Morale	. .	−05	18	07	−11	. .
19. JPQ 9: Independent	. .	. .	04	. .	06	06
20. JPQ 11: Surgency	06	−04	13	. .	. .	−08
21. SSHA Schol. Motivation	10	06	. .	06	28	07
22. CMAS Anxiety	−05	. .	−07	. .	−05	−07
23. CYS Authoritarian	. .	08	03	. .	. .	−05
24. CYS Critic Education	04	. .	−09	−07	06	08
25. CYS Neg. Soc. Orien'n	−11	−05	−10	03	. .	−03
26. CYS Maladjustment	. .	05	. .	03	07	06
27. ISS Family Status	−05	. .	04	. .	. .	. .
28. N-1: Peer Acceptance	25	. .	03	08	07	−12
29. N-2: Model Value	12	. .	−03	05	08	. .
30. N-3: Effectiveness	16	08	. .	. .	08	07
31. N-4: Non-Deviant	05	−10	−08	. .	. .	. .
32. N-5: Quiet Dependency	12	04	14	04	05	03
Multiple correlation, R	75	76	75	68	67	55

Table 5

Orthogonal Normal Varimax Factor Loadings and Communalities of 32 Indicator Variables for 213 Grade VII Students of High Mental Function (IQ 113 to 140)

(Rounded values, decimal points omitted)

Indicators	Factors							h^2
	1	2	3	4	5	6	7	
1. CTMM Mental Function	11	27	52	37	−15	−06	−05	52
2. STEP Listening	21	31	32	23	27	−06	−09	38
3. DAT Mechanical Reasoning	−10	09	59	−03	−11	04	24	44
4. Gestalt Transformation	−01	−01	45	15	−08	17	04	27
5. Rhymes (verbal facility)	−03	13	27	57	17	01	−08	45
6. Unusual Uses	−01	46	18	19	13	06	24	35
7. Consequences	05	56	04	09	−04	−15	01	35
8. Common Situations	05	64	03	−10	−03	−14	−10	46
9. Seeing Problems	00	57	−04	08	05	−01	−01	34
10. Mutilated Words	−03	27	−03	51	−14	−26	05	42
11. Gestalt Completion	05	17	12	22	−07	01	17	13
12. Short Words	−01	−03	−05	64	−03	−02	06	42
13. Dotting	14	17	−04	40	09	−25	−24	33
14. DRT Reaction Time	08	−06	09	35	04	−04	−13	16
15. JPQ 1: Sensitivity	02	04	−29	21	00	03	−55	44
16. JPQ 3: Emotionality	−75	−09	06	−09	06	02	−09	59
17. JPQ 6: Sociability	46	25	−07	10	02	24	−05	35
18. JPQ 8: Social Morale	49	25	−27	15	−18	09	−31	54
19. JPQ 9: Independent	−67	−11	25	−14	−01	01	04	55
20. JPQ 11: Surgency	−09	−02	−15	09	36	−05	−10	18
21. SSHA Schol. Motivation	60	19	05	02	−19	−17	−33	57
22. CMAS Anxiety	−72	14	−19	04	01	21	−02	62
23. CYS Authoritarian	−08	00	01	−12	−35	−03	−09	15
24. CYS Critic Education	−31	−07	04	01	10	21	45	36
25. CYS Neg. Soc. Orient'n	−34	05	−03	08	−08	18	33	27
26. CYS Maladjustment	−56	13	−23	02	09	30	03	48
27. ISS Family Status	11	29	22	−04	12	−06	−19	20
28. N-1: Peer Acceptance	13	21	−11	18	−04	−81	−10	77
29. N-2: Model Value	22	−11	01	07	−62	−01	03	46
30. N-3: Effectiveness	15	19	−11	27	11	−64	−13	58
31. N-4: Non-Deviant	06	−01	06	27	−46	10	−10	31
32. Quiet Dependency	−12	−05	05	11	−13	37	−35	31

Stability and Validity of Factor Variables

The explanatory and predictive values of the factor variables were explored by determining their relative weights for and strength of association with two kinds of criterion measures, academic achievements and age-mate appraisals. Results have been summarized in Table 9 for the general population (N = 740) and for the highly-intelligent HMF students (N = 102) in Table 10, using only those subjects who attended the school from grade seven to grade nine. The "modal weights" shown in the two tables were estimated from the original data (McGuire & Associates, 1960) and show the contribution which a factor variable makes toward the explanation and/or prediction of some kind of valued performance over a specified period of time for a specified population. The loading magnitudes should not be confused with correlation coefficients. They are multipliers or beta weights of sets of factor scores for persons which yield the most efficient explanation and prediction of a specified kind of behavior. In each instance of a significant change in loading magnitudes, from zero to a substantial weight or a shift in signs, two entries have been made in the tables. Thus the weights for factor variables may be viewed as being relative to one another for a given criterion measure, and comparisons can be made across different types of evaluated behavior. Moreover, changes over time due to age and learning experiences or different forms of a criterion measure readily can be identified.

As shown in Tables 9 and 10, the weights attached to the factor variables for each of the criterion measures were different for teacher evaluations, the several kinds of objective measures, and the various appraisals. Nevertheless, within each set for a criterion, the loading magnitudes of a majority of the factor variables were significantly similar from grades seven, to eight, to nine. Boys had a majority of the shifts in beta weights over time and these were confined largely to arithmetic and science in the general population, and to language and science among the potentially-gifted HMF males. The seventh-grade factor variables apparently turned out to be "a much more standard article" (Guilford, 1954, p. 524)

Table 6

Orthogonal Loadings and Optimal Weights of Indicators for HMF Factor Variables

HMF Factor Variables	Indicators of Talented Behavior	Factor Loadings	Optimal Weights	R
1. Personal Stability	JPQ 3: Emotionality vs Ego Strength	−.75*	−.31	.91
	CMAS Anxiety ("to avoid failure")	−.72*	−.27	
	JPQ 9: Independent Dominance	−.67*	−.20	
	SSHA Scholastic Motivation	.60*	.12	
	JPQ 6: Sociability vs Withdrawal	.46*	.12	
	CYS Personal Maladjustment	−.56*	−.14	
	CYS Negative Social Orientation	−.34*	−.08	
	JPQ 8: Socialized Morale	.49*	.05	
2. Divergent Thinking	Common Situations (ideational fluency)	.64*	.33	.79
	Consequences (conceptual flexibility)	.56*	.25	
	Seeing Problems (sensitivity to)	.57*	.24	
	Unusual Uses (spontaneous flexibility)	.46*	.20	
	STEP Listening (awareness to stimuli)	.31	.12	
3. Convergent Thinking	DAT Mechanical Reasoning (eduction)	.59*	.42	.74
	CTMM Mental Function (comprehension)	.52*	.28	
	Gestalt Transformation (redefinition)	.45*	.23	
	STEP Listening (attention)	.32*	.14	
4. Symbol Aptitude	Short Words (recognize symbolic units)	.64*	.40	.80
	Rhymes (word fluency, verbal facility)	.58*	.28	
	Mutilated Words (symbolic closure)	.51*	.19	
	DRT Reaction Time (discrimination)	.35*	.14	
	Dotting (finger and wrist dexterity)	.40*	.13	
	CTMM Mental Function (space relations)	.37	.07	
5. Conforming Behavior	Nom 2: Absence of Negative Model Value	.62*	.46	.75
	CYS Authoritarian Discipline	.35*	.28	
	Nom 4: Nondeviant vs Deviant Behavior	.46*	.22	
	JPQ 11: Surgency vs Desurgency	−.36*	−.12	
6. Age-Mate Acceptance	Nom 1: Peer Acceptance	.81*	.63	.84
	Nom 3: Social Effectiveness	.64*	.19	
	CYS Personal Maladjustment	−.30	−.12	
	Nom 5: Quiet Dependency	−.37*	−.08	
7. Antisocial Attitude	JPQ 1: Sensitivity vs Toughness	−.55*	−.39	.72
	CYS Criticism of Education	.45*	.28	
	Nom 5: Quiet Dependency	−.35	−.25	
	CYS Negative Social Orientation	.33	.15	
	SSHA Scholastic Motivation	−.33	−.09	

* Highest factor loading for the indicator variable in Table I-3; interpretations within parentheses are hypotheses about the factor content of the indicator.

than different sets of indicators with fluctuating combinations such as those shown in Tables 1 and 4.

Another short cut completed our tests of the validity of factor variables since we did not have time for cross-validation studies. This was a test of shrinkage based upon obtained deviations from regression, with appropriate degrees of freedom, when pairs of multiple correlations for the same criterion were compared. In grade seven, the sets of factor scores were not as strongly related to the criteria of talent as were the separate sets of indicators. Instead of shrinking, however, the multiple correlations for the predictions of performances in grade nine either remained similar to or increased significantly over those in grade seven. The only exception was STEP Science for the general population, where forecasting efficiency decreased significantly. For this criterion measure, the coefficient of stability of performance over two forms, 3B and 2A, was relatively low ($r_{I,II} = .45$). Part of the instability was attributed to the fact that many students in the general population did not begin the formal study of a science until their last year of junior high school. In general, the factor variables served to explain a substantial portion of the variation in performances on the STEP instruments and showed improved multiple correlations for the advanced CAT tests of achievement. From the foregoing analyses, we concluded that the factor variables were reasonably valid and stable measures which served to explain a large proportion of the variance in several kinds of talented behavior, both academic and non-academic.

Results and Interpretations

The factor variables developed for the general junior high school population in Table 3, and more fully described in Appendix 1, fell into three categories. Five were common to both sexes and explained a good deal of the variance in different criterion measures. They were *cognitive approach*, *divergent thinking*, *socially-oriented achievement motivation*, *peer stimulus value*, and *absence of negative valuations*. Two were specific for each sex and often served to modify the quality of performance on measures of talented behavior; namely, *anxious emotionality* and *antisocial wariness* for boys,

Table 7

Intercorrelation of Factor Variables for 213 HMF Junior High School Students

(Decimal points omitted)

HMF Factor Variables	Pers. Stab. 1	Diver. Think. 2	Conv. Think. 3	Symb. Aptd. 4	Conf. Behvr. 5	A-M Accept. 6	Anti. Attd. 7
1. Personal Stability	......	1180	−0421	−0086	1210	2385	−3266
2. Divergent Thinking		......	1286	1540	−0106	1606	−0344
3. Convergent Thinking			......	1787	1655	0024	1173
4. Symbol Aptitude				......	−0131	2456	−0880
5. Conforming Behavior					......	−0983	−1841
6. Age-Mate Acceptance						......	−0123
7. Antisocial Attitude							......
Regression on Indicators, R^2	8202	6216	5416	6421	5550	6991	5123
Multiple Correlation, R	9056	7884	7358	8013	7449	8361	7157

and *sensitive dependency* and *perceptual defenses* for girls. The remaining two variables were not linked to either sex role and invariably acted as suppressors to explain less effective behavior; namely, *practical-minded toughness* and *anxious-dependent resentment*.

As shown in Table 9, the ways in which the foregoing factor variables combine vary according to the criterion of talented behavior. Patterns for boys and girls differ in most instances. Two tend to have high relative loadings, *cognitive approach* for standard tests of achievement and *peer stimulus value* for age-mate appraisals. Both are strongly represented in teacher evaluations. The other variables serve to augment, modify, or suppress the ones with major weights.

The HMF factor variables, found in the study of highly-intelligent junior high school students and summarized in Table 6, are not the same as those for the general population: They also are more fully described in Appendix 1. Three are cognitive in nature; namely, *convergent thinking, divergent thinking*, and *symbol aptitude*. Two represent positive and negative elements of the makeup of personality and expressed attitudes, *personal stability* and *antisocial attitude*. The remaining pair combine peer valuations and characteristics of the individual, *age-mate acceptance* and *conforming behavior*.

The seven HMF factor variables serve to explain a significant and substantial amount of the variation in teacher evaluations and in scores achieved on a number of standard tests by the potentially-gifted students. Cognitive abilities, personality attributes, and peer appraisals do not have the same patterns of modal weights from one criterion to another in Table 10. Therefore, one would infer that somewhat different combinations of characteristics are necessary for success in the various subject matter areas, even among potentially able students. Moreover, the patterns of weights on the factor variables for the same measure of talented behavior are not necessarily the same for boys and girls. True, there are some similarities, but one has to accept the conclusion that there are some important sex-typed differences and there are some shifts over time. These differences have an influence upon the manner in which persons of each sex approach and cope

Table 8

Correlation of HMF Factor Variables with Criterion Measures for 213 Students

(Decimal points omitted)

HMF Factor Variables	Criterion Measures					
	Teacher GPA	CAT Reading	CAT Language	CAT Arithmetic	STEP Soc. St.	STEP Science
1. Personal Stability	1658	0020	−0016	0499	1286	0418
2. Divergent Thinking	2667	1989	0463	1355	2257	2828
3. Convergent Thinking	1544	4264	2587	3533	3836	5135
4. Symbol Aptitude	3381	3759	4671	4339	2666	3120
5. Conforming Behavior	1935	−0131	0269	1160	0892	0519
6. Age-Mate Acceptance	4402	2052	2363	1853	1965	1412
7. Antisocial Attitude	−2248	−1049	−2439	−2465	−1253	−0656

with various kinds of valued performances in the junior high school grades.

Conclusion

By and large, we believe that the major outcome of this research is the development of factor variables which fit a model and serve to explain and probably to predict several kinds of talented behavior in the junior high school years. Each set functions somewhat like a differential prediction battery (Horst, 1954). All else being equal, talented behavior is brought about by a combination of innate and acquired abilities, augmented or suppressed by elements of personality and attitudes toward the world of people and things, and facilitated or modified by the expectations and the pressures imposed by peers and older persons. But two other influences are at work. One is the expectations and pressures associated with sex roles and the other involves the context or setting in which learning takes place and behavior occurs. These elements of the prediction model can be represented by factor variables which, so far, have turned out to be stable and valid measures.

Perhaps our greatest concern is with the criterion measures. They do not quite represent what we would regard as the most appropriate criteria of talented behavior, although many are commonly used to evaluate different kinds of performances in school settings. As Michael suggests (1957), there should be some careful exploration of criterion development so that original thinking and desirable kinds of creativity can be evaluated objectively. By and large, the criteria employed in this research place great value upon aptitude with symbols and convergent thinking, the ability to give the one "right answer." With rare exceptions, the data indicate that boys especially are penalized for divergent thinking in the junior high school years. Girls seem to be rewarded more frequently by their teachers when they are more resourceful and less goal-bound. Another concern has to do with the degree to which various forms of conforming behavior appear to be necessary for success in public schools and the extent to which fear and anxiety, on the one hand, and alienation

Table 9

Estimated Modal Weights of Factor Variables for Criterion Measures During Junior High School Years

(Rounded to two places, decimal points omitted)

Factor Variable	GPA Teacher	Criterion Measures CAT Reading	CAT Language	CAT Arithmetic	STEP Soc. St.	STEP Science	Nom Talented	Nom (named)
			Males					(Athletic)
A. Cognitive Approach	37	67	57	49	64	63	03	−07
B. Divergent Thinking	..	02	. .	(−08)(05)†	05	(. .)(09)†	08	−02
C. S-O Achvt. Motivation	15	15	13	06	07	(14)(. .)†	06	−05
D. Peer Stimulus Value	37	10	11	17	06	(06)(. .)†	61	77
E. Abs, Neg. Valuations	19	. .	02	14	05	05	−02	06
F. Anxious Emotionality	. .	(. .)(08)†	. .	(. .)(−08)†	. .	04	05	. .
G. Antisocial Wariness	. .	06	15	(19)(−05)†	03	. .	. .	. .
Range: Gr. VII, VIII, R	.66	.69	.62	.60	.73	.70	.66	.73
Grades IX, R		.79	.70	.68	.69	.64		
			Females					(Rhetoric)
A. Cognitive Approach	41	65	54	53	66	68	13	15

B. Divergent Thinking	13	09	09	05	06	03	−03	..
C. S-O Achvt. Motivation	09	08	(..)(09)†	11	(−03)(11)†	01	..	..
D. Peer Stimulus Value	27	(11)(..)†	15	12	06	05	63	64
E. Abs. Neg. Valuations	17	09	12	18	07	05	−10	09
H. Sensitive Dependency	03	..	..	..	−04	−03	07	07
I. Perceptual Defenses	−08	−05	−05	−04	−06	(−10)(..)†	−03	−04
Range: Gr. VII, VIII, R	.68	.75	.63	.63	.77	.74	.69	.71
Grade IX, R	.73	.78	.72	.69	.71	.65		
Total Population								(IPS)
A. Cognitive Approach	38	67	52	53	53	65	09	−08
B. Divergent Thinking	07	05	08	(−03)(06)†	06	02	..	−06
C. S-O Achvt. Motivation	12	11	(..)(18)†	07	(..)(12)†	(07)(−03)†	05	08
D. Peer Stimulus Value	29	07	11	15	06	(06)(−06)†	62	48
E. Abs. Neg. Valuations	19	05	07	15	05	05	−06	23
J. Prac-Mind. Toughness	−13‡	−14‡	−35‡	−17‡	−12‡	(−08)(..)‡	−02	
K. Anx-Depnt. Resentment	−06‡	−04‡	−07‡	..	−05‡	..	03	
Range: Gr. VII, VIII, R	.69	.73	.73	.62	.74	.71	.66	.50
Grade IX, R	.71	.77	.68	.70	.69	.66		

---

† Loading weights for the factor variable changed from grades VII and VIII, first (), to grade IX, second (), indicating a shift in the contribution the factor scores made toward prediction of the criterion variable.

‡ Loading weights for factor variables H and I were obtained for the total population only in grades VII and VIII. Regressions of criteria on these factor scores were not computed for grade IX, nor for IPS Peer Status.

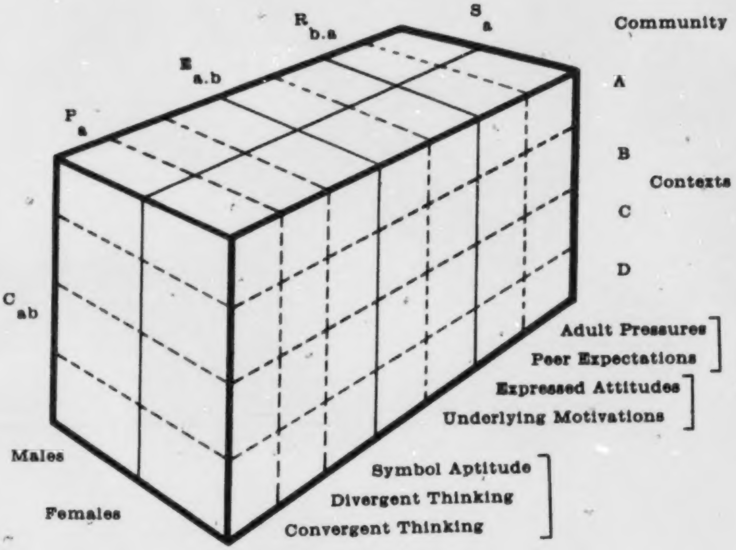

$R_{b.a}$
S_a
Community
$E_{a.b}$
A
P_a
B
Contexts
C
D
C_{ab}
Adult Pressures
Peer Expectations
Expressed Attitudes
Underlying Motivations
Males
Symbol Aptitude
Divergent Thinking
Females
Convergent Thinking

Fig. 1. Schematic diagram of a model for research in human talent.

$$B_a = f(P_a, E_{a.b}, R_{b.a}), S_a, C_{ab}$$

B_a = behavior of the person (a) to be explained or predicted;

P_a = potential cognitive, perceptual, and other relevant abilities;

$E_{a.b}$ = elements of personality and motivation, especially expectations about one's own behavior and probable responses of other persons;

$R_{b.a}$ = responses of other persons (b) expressed in terms of their expectations and pressures they impose upon the given individual (a).

S_a = sex-role identification of the individual (a) and sex-typing of socialization pressures, both moderating preceding variables;

C_{ab} = context of behavior, such as a community or school setting which provides an institutional framework along with certain experiences and impersonal expectations; or, the setting in which a natural or a laboratory experiment takes place.

expressed in antisocial attitudes, on the other, almost invariably suppress the quality of observed performances.

On the positive side, the research points to the eventual development of a battery of factor variables to assess the present status and to forecast future performances in secondary schools and, eventually, of college students. For this purpose, the factor variables emerging from the study of the highly intelligent students seem to show great promise. A small battery of instruments could be devised to represent the various abilities, attributes of, and pressures imposed upon boys and girls. They could be used to identify persons who could benefit from special attention either in the sense of carefully devised learning experiences or from counseling. In addition, such a battery could be employed to predict different kinds of performances if no provision is made for what we have termed educational telesis.

Table 10

Estimated Modal Weights of HMF Factor Variables for Criterion Measures During Junior High School Years for a Population of Highly-Intelligent Students from Four Communities*

(Rounded to two places, decimal points omitted)

HMF Factor Variables	Criterion Measures						
	GPA Teacher	CAT Reading	CAT Language	CAT Arithmetic	STEP Soc. St.	STEP Science	STEP Math
Males							
1. Personal Stability	−09	−13	(−13)(. .)†	(−12)(02)†	−09	(17)(−14)†	23
2. Divergent Thinking	08	−03	(. .)(−12)†	(21)(−03)†	10	(−03)(13)†	06
3. Convergent Thinking	12	35	21	15	23	36	47
4. Symbol Aptitude	. .	25	47	29	06	(−09)(21)†	13
5. Conforming Behavior	26	(14)(. .)†	(15)(. .)†	15	(09)(−14)†	12	07
6. Age-Mate Acceptance	66	20	13	29	26	06	12
7. Antisocial Attitude	−09	−22	−15	−19	−19	−21	−03
Multiple R, Grade VIII	.67	.56	.61	.63	.62	.53	
Grade IX		.66	.69	.62	.60	.60	.60
Females							
1. Personal Stability	14	05	13	11	08	06	11
2. Divergent Thinking	15	17	(−06)(14)†	16	19	(24)(. .)	16

3. Convergent Thinking	07	35	..	27	42	39	25
4. Symbol Aptitude	39	27	57	28	11	26	34
5. Conforming Behavior	09	09	−13	20	10	14	..
6. Age-Mate Acceptance	27	13	13	15	05	(−13)(..)†	09
7. Antisocial Attitude	−22	−09	−23	−16	−08	..	..
Multiple R, Grade VIII	.66	.60	.66	.66	.61	.58	
Grade IX		.69	.70	.65	.56	.61	.60
Total Population							
1. Personal Stability	..	−04	(−11)(03)†	−03	..	(11)(−03)†	17
2. Divergent Thinking	12	10	..	11	16	(20)(..)†	10
3. Convergent Thinking	06	37	18	21	40	44	45
4. Symbol Aptitude	24	27	49	34	12	13	19
5. Conforming Behavior	17	..	..	11	06	(−03)(12)†	06
6. Age-Mate Acceptance	38	15	13	20	14	04	08
7. Antisocial Attitude	−20	−16	−23	−20	−14	−07	05
Multiple R, Grade VII	.60	.57	.58	.58	.50	.60	
Grade VIII	.62	.56	.62	.65	.61	.49	
Grade IX		.64	.68	.61	.56	.62	.60

---

* Table 1 was derived from Tables I-3, I-9, and I-10 in Appendix I. GPA Teacher Evaluations were available only for grades VII and VIII STEP Mathematics was substituted in Grade IX. Regressions on criterion measures were computed separately for boys and girls only for grades VIII and IX, 53 males and 49 females. The original population of HMF students in grade VII was 213, of whom 120 were males and 93 females.

† Loading weights for the factor variable changed from grade VII or VIII, first (), to grade IX, second (), indicating a shift in the contribution of the factor score toward the prediction of the criterion measure.

Appendix 1

Description of Factor Variables

TWO SETS of factor variables were developed in the study of talented behavior in junior high schools. One set emerged from the analyses of data gathered on 1,242 students, 634 boys and 608 girls, of four Texas communities beginning in the seventh grade and followed through the ninth grade where only 730 were still at the same school location. The other set was derived from data on 213 highly-intelligent (HMF) or potentially-gifted adolescents, 120 males and 93 females, from the same high school populations.

FACTOR VARIABLES FOR JUNIOR HIGH SCHOOL STUDENTS

A set of eleven factor variables served to explain and predict the various kinds of talented behavior observed among all junior high school students. Factor-analytic techniques were used to sort out and map the several kinds of abilities and other attributes. Then, for each of the factor variables, a multiple regression procedure was employed to estimate the optimal weights of each of the indicators of talented behavior.

A. *Cognitive Approach*. Efficiency in attending to and concentrating upon verbal stimuli during the learning process (STEP Listening); recognition of appropriate answers to problems involving spatial relations, verbal concepts, logical and numerical reasoning (CTMM); ability to shift the function of a part of an object and use it in a new way (Gestalt Transformation); capacity to visualize and formulate complex conceptions from figural materials (DAT Mechanical Reasoning); five other indicators have much smaller weights.

B. *Divergent Thinking*. Ability to think in different, less goalbound directions demonstrated by going beyond what is given and extrapolating outcomes (Consequences), awareness that problems exist (Seeing Problems), and calling up as many ideas or responses as possible in a given time (Common Situations).

C. *Socially-Oriented Achievement Motivation*. Acceptance of school and cultural standards (JPQ 8); scholastic motivation, accompanied by effective study habits and attitudes (SSHA); representing oneself as a stable, confident person (JPQ 3); four additional indicators.

D. *Peer Stimulus Value*. Acceptance by one's peers (party with, like to be like, work with on a school problem); expected to be an effectively-functioning person (enjoys everything, self confident, shows initiative); seldom regarded as a quiet, dependent individual.

E. *Absence of Negative Valuations*. Not necessarily recognized or accepted by age-mates but seldom named as a negative model (not like to be like, not ask for help, not party with); seldom nominated by peers for deviant behavior (dislikes school, has to be told, gets by).

Factor-matching techniques showed that the five foregoing factor variables were common to both boys and girls. The next four were specific to males and females and two were identified for each sex.

F. *Anxious Emotionality (Boys).* Self reports indicate "a motive to avoid failure," especially in ego-involving, threatening, or stressful situations (CMAS Anxiety); misperception, inability to cope with pressures, and lack of a sense of identity (CYS Maladjustment); emotional sensitivity (JPQ 1); reacts emotionally (JPQ 3).

G. *Antisocial Wariness (Boys).* A set to accept authoritarian attitudes and awareness of controls exerted by authority figures (CYS Authoritarian); somewhat alienated from others (CYS Negative Social Orientation); negative toward teachers and critical of what is expected in school (CYS Criticism of Education); perceives symbolic materials quickly, even from partial clues (Short Words, Mutilated Words).

H. *Sensitive Dependency (Girls).* Expresses authoritarian attitudes and accepts control of authority figures (CYS Authoritarian); emotionally sensitive, tender-minded (JPQ 1); complies with most school and cultural standards (JPQ 8); quiet, serious (JPQ 11).

I. *Perceptual Defenses (Girls).* Originally named "Perceptual-Motor Skills" but reinterpreted; identifies objects with parts missing (Gestalt Completion); careful to represent self positively (CYS Neg. Social Orientation); quick and careful reactions (Dotting, DRT Reaction Time).

The remaining two factor variables were identified in the analysis indicators for the total population. Neither of them was correlated with factors obtained for boys or girls separately, so the attributes which they combine could not be linked to either sex role. They were regarded as moderator variables which combined cognitive and noncognitive measures (J) or elements of personality and expressed attitudes (K).

J. *Practical-Minded Toughness.* Represents self as tough-minded and scoffing at emotions (JPQ 1); probably has mechanical experience and responds to practical problems (DAT Mechanical); lacks the ability to perceive and cope with symbolic materials (Mutilated Words); low speed of closure on word units (Short Words).

K. *Anxious-Dependent Resentment.* Self reports indicate underlying anxiety (CMAS); expresses authoritarian beliefs (CYS); distrustful, egocentric, and resentful (CYS Neg. Social Orientation); desurgent, passively awaits what is going to happen (JPQ 11).

HMF FACTOR VARIABLES FOR POTENTIALLY-ABLE STUDENTS

1. *Personal Stability.* High stability or ego strength (JPQ 3); low fear of failure (CMAS Anxiety); cooperative, restrained, and able to communicate (JPQ 9); five additional indicators with lesser weights.

2. *Divergent Thinking.* Ability to think in different, less goalbound directions marked by ideational fluency (Common Situations); conceptual flexibility (Consequences); sensitivity to, or awareness that problems exist (Seeing Problems); some originality, or spontaneous flexibility (Unusual Uses).

3. *Convergent Thinking.* Ability to channel thinking to one unique conclusion by means of eduction (DAT Mechanical Reasoning); comprehension (CTMM); redefinition (Gestalt Transformation); and attention to stimuli of a verbal nature (STEP Listening).

4. *Symbol Aptitude*. Rapid recognition of symbolic units (Short Words); facility with verbal materials, word fluency (Rhymes); perception of symbols when given only partial clues, symbolic closure (Mutilated Words); three additional indicators with lesser weights.

5. *Conforming Behavior*. Absence of negative valuations by agemates; acceptance of authoritarian beliefs and awareness of controls by authority figures (CYS Authoritarian); seldom regarded as deviant in behavior by peers.

6. *Age-Mate Acceptance*. Acceptance by one's peers as a person to be like or be with; expected to be an effective, mentally healthy person; able to cope with pressures and possesses a sense of identity (CYS Personal Maladjustment).

7. *Antisocial Attitude*. Toughminded outlook (JPQ 1); critical of school and teachers (CYS Criticism of Education); seldom regarded as a quietly dependent individual; somewhat alienated from other people (CYS Neg. Social Orientation).

REFERENCES

Atkinson, J. W. Motivational determinants of risk-taking behavior. *Psychol. Rev.*, 1957, 64, 359-372.

Bloom, B. S. (Ed.) *Taxonomy of educational objectives*. New York: Longmans, Green, 1956.

Cattell, R. B. *Personality and motivation structure and measurement*. New York: World Book, 1957.

Elkins, Deborah. Some factors related to the choice-status of ninety eighth grade children in a school society. *Genet. Psychol. Monogr.*, 1958, 58, 207-272.

Ferguson, G. A. On learning and human ability. *Canad. J. Psychol.*, 1954, 8, 95-112.

Ferguson, G. A. On transfer and the abilities of man. *Canad. J. Psychol.*, 1956, 10, 121-131.

Goethals, G. W. A framework for educational research. *Harv. educ. Rev.*, 1958, 28, 29-43.

Greenberger, M. H., & Ward, J. H., Jr. An iterative technique for multiple correlation analysis. *IBM Technical News Letter*, 1956, No. 12, 85-97.

Guilford, J. P. *Psychometric methods*. (2nd Ed.) New York: McGraw-Hill, 1954.

Guilford, J. P. The structure of the intellect. *Psychol. Rev.*, 1956, 53, 267-293.

Guilford, J. P. Three faces of intellect. *Amer. Psychologist*, 1959, 14, 469-479.

Haggard, E. A. Socialization, personality, and academic achievement in gifted children. *Sch. Rev.*, 1957, 65, 388-414.

Hindsman, E., & Duke, R. L. Development and utilization of talent. *J. teach. Educ.*, 1960, 9, 107-116.

Horst, P. A technique for the development of a differential prediction battery. *Psychol. Monogr.*, 1954, 68, No. 9 (Whole No. 380).

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, 23, 187-200.

McGuire, C. The Textown study of adolescence. *Texas J. Sci.*, 1956, 8, 264-274.

McGuire, C., & Associates. *Talented behavior in junior high schools.* Final Report, Project No. 025, Cooperative Research Program of the U. S. Office of Education, Department of Health, Education, and Welfare. Austin, Texas: University of Texas, 1960. (Copies distributed to certain libraries by the U. S. Office of Education.)

Michael, W. B. Differential testing of high-level personnel. *Educ. psychol. Measmt.*, 1957, 17, 475-490.

More, D. M. Developmental concordance and discordance during puberty and early adolescence. *Monogr. Soc. Res. Child Develpm.*, 1953, 18, No. 1 (Serial No. 56).

Sears, R. B. A theoretical framework for personality and social behavior. *Amer. Psychologist,* 1951, 6, 476-484.

Stern, G. G., Stein, M. L., & Bloom, B. S. *Methods in personality assessment.* Glencoe, Ill.: Free Press, 1956.

Wilson, R. C. Creativity. *In Education for the gifted in school and college.* 57th Yrbk., Nat. Soc. Stud. Educ., Part II. Chicago: Univer. of Chicago Press, 1958. Pp. 108-126.

Luncheon Address

Arthur S. Adams,
President, American
Council on
Education

The Pace of Change

LADIES AND GENTLEMEN, I feel very brash indeed in appearing before you today. As some of you know very well, I had the privilege of standing on this very rostrum just the day before yesterday. I enjoyed the experience, but my resources are not such that I can successfully emulate the activities of our presidential candidates and speak daily on a variety of subjects.

What I have to say this afternoon is known to all of you. There is nothing new about it. I propose to factor no matrices. I shall compute no coefficients of stability of performance, but I do applaud those who can and do such things and obtain the facts upon which to base their conclusions, for surely we make progress only in the measure in which we determine the facts solidly and then proceed to reason to sensible conclusions.

Unhappily, since education is a subject which has preoccupied the mind of man for centuries, probably ever since the cave man era, we have many theories of education. Some are spun out of the gossamer of imagination and prejudice. They are always given wide distribution in the public press. These theories are not based upon facts, they are based upon supposition and prejudice and we who work in education have to keep our blood pressures down, continue to ascertain facts and to think with the very best zeal we can muster in order to deal adequately with these proposals.

What causes this continuing and sometimes raucous debate about educational policy and practice? In my judgment it is inherent in the mind of man to give great concern not only to the propagation of the race, but also to its education. I have noted this in remote places in the earth—in East Africa, for instance, where the average gross national product per individual is $14 a year, $14 a year per *individual*—I have noted this concern in their attitude toward education. In the educational system available to them, the parents must put up the equivalent of five pounds, which is about $14. One wonders what a family does for subsistence since, on the average, the whole part of the national income accruing to the individual is represented in educational expenditure.

I have visited crude schools: mud and wattle huts covered with thatched palms, no doors, no windows, no floors, simply rude benches and a teacher standing before the class with a dog-eared book in his hand with forty or fifty African children sitting there with their hands quietly folded in their laps and their eyes glued on the teacher (no wiggling in these schools). During every such visit, I have had the feeling that if we could have the measure of commitment to education in our country that they make so manifest, we could move ahead very rapidly indeed.

Let me inject another note just to show you what this dedication amounts to: I talked with a teacher of what they call the fourth form in that country, a very intelligent, capable African. I said to him, "What one thing do you think would give you the greatest help in carrying on your work?"

I thought perhaps he might say a mimeograph machine or a projector or some of the other things that one would think natural in this country. His reply was, "If I could only have a blackboard!"

This, then, is what we have to think about as we consider our commitment to education. Here are people who in the last generation were primitive pagans and who are now committed to education to a far greater degree, I believe, than we are as a total people.

My purpose, as I said at the outset, is not to give you any new, brilliant concoction that has just come out of my head. It is only to remind you of things that you already know, but which, in the press of affairs, may not be at the center of your thoughts.

Everyone knows that there is something about change or motion that always captivates human attention. The child amusing himself with the changing designs in the kaleidoscope, the advertising displays we see in the store windows as we walk up and down the streets of this great city attract us because in them something moves. Even perhaps the eagerness with which some people go for automobile drives is an indication of our inherent desire to move.

I came across a statistic which was revealing to me in this connection two years ago when our Post Office Department in Washington reported that in the previous year 40 million Americans had changed their address. This may be a manifestation of that old saw that it is cheaper to move than to pay rent. I don't know; I don't think it is. I think there is a certain restlessness which motivates all of us, and perhaps in our country this restlessness reaches a zenith. We do like change and at the very same time, ladies and gentlemen, we resist it and this is one of the classic paradoxes of human behavior.

For example, the college president who sees that the institution might profitably undertake a different course of action had better not announce this change if he values his hide and his position. First he brings in a group of senior deans and, after a casual and genial discussion of the state of the weather, the performance of the football team and what the chances are for getting a little more money from private donors or the legislature, he says, "Oh, by the way, I was thinking the other day about a certain matter and discussing it with a friend and he suggested" (never the president, mind you) "he suggested that we might do thus and so."

And the matter is discussed, and he says, "You might try it out on some of the staff members and see what they think of it." The deans then go to the appropriate faculty members and, in this same ingenious way, the idea is gently introduced.

Finally, the idea turns out to have some acceptance and merit. Then it is debated at length and perhaps fifteen, eighteen or twenty-four months after the president has first operated by induced initiative to plant the idea in somebody else's mind, it comes as a demand from the faculty that this be done. This only happens if the president is very wise and fortunate. At this point, then, the change is made and it may have been a very simple matter indeed.

Now you all reveal by your reaction to my comments that you understand this resistance to change in education. Yet at the same time—at the very same time—ladies and gentlemen, the pace of change in our society is constantly accelerating.

Let me detail three aspects of that pace of change. The first has to do with the pace of change in science and technology.

I am sure that I need not dwell on these for I am confident that this audience is well aware of the impact that changes in transportation and communication have brought about in our society. We hardly think about it when we travel at a speed of ten miles a minute, six miles up. This is casual. We accept that; and yet when we come to education, we take a diametrically opposite position and firmly hold to the position that not a thing shall be changed.

Let me recall to your mind that our current complete acceptance of television as a fact in our social life today has been brought about in just 15 years. Similarly, the advance in nuclear technology, which has such terrifying implications, had its start less than a generation ago.

When the late Professor Ernest O. Lawrence, then an Associate Professor, first came to the University of California in 1928, he announced what seemed then to be the preposterous notion that he was going to devise some means of collecting isotopes of the various elements in discretely measurable quantities. Previously, the existence of isotopes of the elements had been detected in band spectra, but they occurred in such minute amounts that no one conceived that anybody could get enough of an isotope to weigh even a milligram. I was there. I regret to say that I participated in luncheon discussions in which, with biting sarcasm and irony, those present chided Professor Lawrence on the ridiculousness of his notion.

When he sought to obtain money to build the first accelerator, he had even greater difficulties, but Professor Lawrence was not to be denied. He visited every junk yard within fifty miles of Berkeley, California, and from all sorts of discarded machinery he was able to find in those junk yards he assembled a most curious looking contraption which was the first electron accelerator. And that was only a little more than thirty years ago! Today we have not only fairly good understanding, but also application, of both the principles of atomic fission and atomic fusion with enormous releases of energy resulting therefrom.

This is the pace of change.

Then, too, the whole electronic industry has come of age in little more than a generation. In another area, there is no need to detail the tremendous role in our economy played by the manufacturers of all sorts of labor-saving equipment and gadgets for the home. I presume that no home will be complete in the next few months without an electric can opener. One has only to visit a model home in some new subdivision to see at once why equipping the home of today accounts for very nearly half of its total cost. We have only to look around us, if we can find a moment to do so, to see quickly the enormous magnitude of the effect that technical advance has had on the lives of everyone of us. And we have adjusted to those changes.

There is no doubt either that, except for a national emergency, this advance will continue at an accelerated pace. If there is a national emergency, it will probably accelerate anyway, but in a far different direction. Thousands of minds are busily occupied today in devising new ways of doing old things. Other thousands of minds—indeed, our best minds—probe relentlessly to extend the frontiers of science and technology, frontiers which are no longer limited to the dimensions of this earth but reach out into interstellar space. And again we seem to be able to adjust to that without too much trouble. We know all this if we but stop to realize it. The pace of change is far more rapid in this area than we generally realize and it is likely to increase.

My question is this: What does this do to our outlook, to our understanding of our society and of the kind of world in which we live?

We cannot decry technical advance, for the mind of man in its relentless quest for more and more knowledge will not be denied. We cannot turn off technical advance as we would turn off the water spigot in the kitchen sink. What are we to do then? We must seek to understand, to reorganize what we know as well as cur outlook upon it. We must seek ways in all of this complex and accelerating change by which the human being may be the master and not the servant of technical change.

Yet, this pace of change that has resulted from scientific and technical change is by no means the only change with which we must reckon. The sheer magnitude of present-day problems which society is called upon to solve requires that new approaches be devised for dealing with those problems. It is not enough that these approaches merely be new, although there are some who would argue in this fashion. The approaches must be relevant. They must take into account all of the factors involved and they must be calculated to lead to useful results.

In education, just in the past few years, we have heard a great hue and cry raised about the inadequacies of American education, with all sorts of proposals advanced to remedy alleged limitation. Fundamentally, it seems to me that a factor which we have not yet fully considered has to do with the numbers of individuals seeking education at all levels. Oh, yes, we have all heard about the population explosion—the rising birth rate, the tidal wave of students and all the rest—but have we recognized that these dramatically increased and increasing numbers really pose a new and a different kind of problem, not simply an extension of the problem that we knew before?

Again, let us look at some of the advances in physics. The whole business of releasing atomic energy depends, in the last analysis, on the critical mass of the materials brought together in a reactor. It is a matter of critical mass: if that mass is not present, nothing happens. If that mass is present, something happens with enormous effect. I suspect there is something like a critical mass in educational problems, and I would say that we'd better think of the dramatically increased numbers with which we must deal in that sense.

I am confident that in all this the principles of American philosophy with respect to education remain valid and, moreover, are held with conviction by all of the American people. The principles of equality of educational opportunity, of decentralization of educational authority, and of humanitarian concern for the individual seem to me to be at the very core of what our people really believe about education. As we tackle the massive problem of numbers, I am certain that we shall be successful only to the degree that we adhere firmly to these principles.

It would be easy to devise new educational procedures based upon other principles. Indeed, we know that the Soviet Union does just that. The individual has no significance in the Soviet system. He is merely a pawn of the state, a naked statistic totally lacking in human identity. Our approach to this massive problem of numbers must be made in terms of principles we believe to be right, principles which, in the end, are the ones which every emerging nation in the world seeks to emulate.

But how shall we do this? Does this mean that we simply keep on doing what we have done in the past? Only more so? I think not.

I am confident that we must seek new and better ways of matching the individual to the opportunity for an education that will best provide means for developing his talent and his interest to the maximum. The Educational Testing Service and other similar agencies have long been concerned with this very problem and have used all of the technical advances currently available to do more effective work in this regard. Yet, there is a caution here, too. With the pace of technical advance having reached the point where from three to six thousand test papers can be scanned by a machine in an hour, one could become beguiled with the marvelous efficiency of such a process and at least momentarily forget the hopes, the aspirations and the future of the individual human being tied up in every one of those papers.

I say this because I myself have watched this scoring machine and others like it; and I have tended to become fascinated with the flashing lights, the rapidly moving papers and the whirr of machinery. There is something in my nature that just responds to that sort of stimulation, and it becomes a genuine effort for me to focus my mind on the fact that every one of those papers represents one step, and perhaps a crucial one, in the chance for growth of an individual boy or girl.

Is the worth of a human being to be described uniquely by an item on an IBM card? Of course not. I am fully convinced that this was never the motive nor is it the motive of those concerned with the whole process of testing. The true motive is to use all of the techniques available in order to help place the individual where he can realize his maximum potential. But, again, I say, such is the fascination of the human being with change, no matter of what sort, that this motive may not be understood by the very people who give the greatest credence to the results.

I have chosen education as my illustration for the pace of change in society because of its natural relevance here today. Yet I would certainly be remiss did I not at least mention a number of other factors of social change which surely have bearing on the subject.

The distribution of population by age group is one such factor. Because of the rise in the birth rate on one hand and the advance of medical science on the other, our population today has nearly 75 per cent more people under 20 and over 65 than it had in 1940. Yet the number of people between 20 and 35 is only 3 per cent greater than in 1940 and the number in the age groups between 35 and 60 is only about a third greater than twenty years ago. These facts mean that education for the young and care for the aged have assumed new and greatly increased importance to our society.

We surely have evidence of this in the Congress and in the current presidential campaign. Again we have problems to solve of a magnitude that we never dreamed of before and of necessity this magnitude must change our approach appreciably. A mere extension of what we did before will not do.

In both of these instances, it is not only a question of exerting greater effort or of exhibiting greater concern, it is a question of providing the money necessary to support the effort. This, in turn, is bound to have its effect on our tax structure and, in the end, on our economy as a whole. Yet the problem must be solved. It will never be solved if it is dealt with in splinters.

I submit that the only way to solve these problems satisfactorily is to recognize the full magnitude which they really have; to gather and analyze all of the facts relating to them and then to seek to arrive at reasonable conclusions, rather than emotional ones.

A third area in which the pace of change is breathtaking has to do with the relationships of our country with the other countries of the world. In trying to appraise this rate of change I have adopted a personal appraisal factor, which is that events in the world are happening just about ten times faster than I ever thought possible.

Just three years ago, I had the privilege of visiting the Belgian Congo which I thought then was the most tranquil African state I had seen. I attended several meetings of Africans at which the question of independence for the Congo was discussed. I listened to one speech by an African who said he thought the Congo was in no hurry for independence and if it got it in thirty-five to fifty years this would be all right with them. That was just three years ago. Do you see where I get my factor of ten?

We now know that while the Congo is no longer under Belgian control (indeed, it is hardly under any control), it can scarcely be said that it has achieved responsible independence.

Now, on a brighter side, I can report that other African countries with which I am familiar have made changes in their educational systems that are almost unbelievable in this same period of three years. Indeed, my factor of ten still applies here, for the advance has certainly been ten times faster than I would have thought possible. Others who have traveled more widely and have visited various other emerging countries of the world tell me that they have witnessed this same sort of truly astonishing change.

These considerations, too, have great meaning for education. No longer can we be satisfied with an education exclusively oriented to our traditional western point of view. In this connection, let me relate an anecdote that occurred the other day. At the Washington International Center, which is operated by the American Council on Education and where we give a one-week introduction to American life, I was talking with an Indonesian, a very bright young chap with a twinkle in his eye. In my conversation with him, I used the word "oriented"—a natural enough thing, a word which we use all the time, without being quite sure what it means.

At this point, my young Indonesian friend looked at me with a smile. He said, "But, sir, I don't wish to become oriented. I want to become Occidized."

No longer can we think that the customs, the philosophy, or the objectives of the peoples of the far corners of the globe from us are matters that can be ignored or, worse yet, can only be thought of as divertingly bizarre. All of these peoples are now our neighbors and what happens to them is of immediate and pressing concern to us. In the years immediately ahead, as is so well demonstrated in the membership roll of the United Nations, we must have knowledge of these nations and we must cultivate genuine understanding of their ways and of their aspirations.

This is yet another of these massive problems of change, and I could not, I *would* not, presume to offer any pat solution for it. But I am convinced that the way to start is to recognize it as a massive problem and then to seek all possible knowledge in relation to it in that context.

On this first trip I made to Africa three years ago, if you will permit me to become personal again, I confess unashamedly that every notion I had about the African peoples as I set out on my trip proved to be wrong when I got there. I can hardly remember now what those notions were, so great was the impact of reality upon me. Yet I certainly thought I would find the Africans lethargic, easy-going, careless and somewhat dull. The facts turned out to be exactly the contrary. I found them alert, vigorous and decidedly demanding. They loved to talk and they loved to talk far into the night. Yet they always set my first appointment in the morning at eight o'clock, irrespective of whether I got into bed at two, three or four in the morning.

Now, it may well be that I was simply insensitive to what I should have known about Africa. If so, I confess it. Yet, the fact remains that virtually every traveler with whom I have talked reports precisely the same experience: a complete change in his personal orientation to the notions that he had had before.

This is why I argue so vigorously for the seeking of greater knowledge of all of the countries of the world, especially in those areas of the world in which new nations are emerging or have recently emerged. We in the United States just don't know enough about these areas to be able to come to usefully valid conclusions about our relationships to them.

Now, ladies and gentlemen, I have sought to identify some of the highlights of what I conceive to be the pace of change that results from rapidly accelerating developments in science and technology, from the new magnitude of social problems we face, and from new world conditions. All of these have immediate, direct bearing upon education.

In all of them, we can be assured that the pace is rapidly increasing. We can no longer be satisfied simply by doing more energetically what we have done before. In so doing, we are apt only to increase activity, activity which confuses our minds, saps our energy and beclouds our perspective.

As I see it, our great need today is to come to a fuller realization of this pace of change. No matter what the seeming demands of the moment, no matter how frantically we try to keep up with the distracting demands that the complexities of modern society put upon us, we must recognize that it is the height of futility to confuse activity with achievement. This is the meaning of the pace of change.

Session II

Remarks of the Chairman

THE THEME for this afternoon is "Testing in the Language Arts." I recognize that the phrase "language arts" may have a sour taste for some. The Council on Basic Education appears to feel that this phrase is a nice example of "pedaguese," and that courses in the teaching of "language arts" are intellectual sumps of a low order, about on the level of driver training and safety education. I will grant that some things which pass as "language arts" have hardly anything to do with language and are far from artistic. Nevertheless, I find this is a useful and innocuous phrase for referring collectively to such things as courses in English language and literature, foreign languages and literatures, and yes, even public speaking, oral interpretation and forensics. The phrase "language arts" commits us to no special point of view, in my mind at least.

This afternoon's theme brings us to an emphatically interdisciplinary part of the conference—not that there are any parts of education which are *not* interdisciplinary (and I sometimes think that its interdisciplinary nature is at once the virtue and the weakness of the enterprise of education). But this afternoon we are reaching out to disciplines other than psychology and statistics. The first two of our speakers are, by profession and training, in the field of the liberal arts; in their work in educational measurement, they are likely to be cast in the role of what test makers sometimes call "subject matter experts." Let us remind ourselves of the respect we must accord to the subject matter expert—he must be an equal partner—indeed, *primus inter pares*, "first among equals"—in the building of tests. It happens, however, that our speakers from the liberal arts this afternoon can be even "more equal," because both of them have had sufficient experience in test making to have become much more thoroughly acquainted with testing problems than many, I suspect, of their colleagues.

The first speaker will be Dr. Harold C. Martin, Director of General Education A at Harvard, that is, the course in Freshman English composition. Dr. Martin took his doctorate in comparative literature and has been a high school teacher of English and of French, and later a high school principal, before he became associated with Harvard. Recently he was appointed to be Chairman of the College Entrance Examination Board Commission on English, established to help in improving the teaching of English in our secondary schools. His topic is "Testing for Elegance."

Our second speaker, Dr. Wilmarth H. Starr, formerly of the University of Maine and currently with the Modern Language Association of America before he takes up his duties at New York University next year, will introduce to us some new developments in the field of foreign language testing. It is in connection with his talk that we can pay our respects to the National Defense Education Act, as I mentioned this morning, since one of the provisions of that Act is for studies of methods and materials in the teaching of foreign languages. Dr. Starr is the director of a project to develop new tests of competency for teachers of foreign languages; this project is being conducted by the Modern Language Association under contract with the Language Development Section of the U. S. Office of Education. When he is not working on projects like this, he is a professor of French language and literature. His topic is "Competency First: New Tests in Foreign Languages."

Our third speaker will be an old friend to many of us, in fact a former chairman of these Invitational Conferences, Dr. Irving Lorge of Teachers College, Columbia. Dr. Lorge has spent most of his professional life working on problems of teaching the language arts; the Lorge readability formula, the Lorge semantic count, the Lorge-Thorndike Intelligence Test, and a number of other things have made *Lorge* almost a household word among educational measurement specialists and educators generally. He will speak about "Estimating Structure in Prose."

Harold C. Martin,
*General
Education A,
Harvard University*

Testing for Elegance

WE ALL KNOW that there is widespread discontent among laymen at even the best tests so far devised for measuring students' ability to write English prose. That discontent forces reconsideration on any who are concerned to protect the integrity of measurement as well as the welfare of students. Such reconsideration goes on constantly among testing experts, as I am well aware, and it may be impertinent for a layman to presume in a field so recondite. Yet I am moved to do so because I believe that a puristic approach to the problems such testing raises does more harm than good at this time and because, in at least one instance, the lay voice has prevailed with a result that seems to me intrinsically bad.

My formal acquaintance with testing is limited to seven years on the Committee of Examiners in English of the College Board and to less extensive work on the Advanced Placement Examinations of the Board in their early years. Now the English Composition Test and the Advanced Placement Examinations in literature and composition may not be the best so far devised, but they have at least attempted something beyond the measurement of rudimentary skills (capitalization, punctuation, spelling) and of elementary information (common usage, parts of speech, identification of syntactic structures and so on). Because the English Composition Test is designed for massive administrations requiring quick correction and report, it has, as you probably know, developed oblique procedures. It purports, for instance, to measure the ability to organize a paragraph or a complete essay by requiring the examinee to put scrambled sentences in logical, topical, or chronological order, or to create from undifferentiated items a structure of principal, subordinated, and coordinated parts.

On occasion it has used one original line of a quatrain along with three spurious lines to measure the examinee's sense of meaning, his ability to distinguish appropriate from inappropriate diction, and his ear for linguistic rhythms. Somewhat more directly, it has set him tasks of re-writing sullied prose as a way of testing his capacity to compose felicitous prose of his own.

I name here only the three most successful devices tried in the English Composition Test. My files show at least a dozen other devices pre-tested with varying amounts of success but seldom with enough to encourage use in a large-scale administration.

Those approaches used and re-used in recent years on the test have had reasonable success as success is measured in such matters. That is, they have shown themselves to be—either taken alone or in conjunction with other tests—decently satisfactory predictors of the ability to write such prose as the college freshman year demands insofar as grades at the end of the freshman year are themselves reliable measures of competence. To be precise, the ECT is less successful than the Scholastic Aptitude Test—Verbal as a predictor but, taken in conjunction with the SAT-V, is more successful than either of the two used alone.

Perhaps I should qualify that definition of "success" to allow for a statistical subtlety I have so far not been able to understand. In the analysis of results on a multi-section test it is customary to congratulate oneself if each part score shows a strong similarity to the whole score; that I think, is called high reliability of parts. On the other hand, I have heard sharp deviation of one score from other scores on the same test approved as evidence that the deviating part is successful because it measures something different from what the other parts measure. There may be more compatibility in the two positions than I am aware of; in my present ignorance I must confess that the apparent incongruity makes me uneasy. Despite the success, however defined, of the ECT test, the outcry against it has grown rather than diminished in the past ten years. Last year it found vent in a move to substitute a direct measurement of writing ability by the simplest and most obvious procedure of all, that of requiring students to write an essay on an assigned topic.

The attractive simplicity and common sense of such a procedure proved, when the chips were down, more powerful than any arguments that could be brought against it, and that is not to be wondered at. But the new test—or "writing sample" as it is now euphemistically called—is, I believe, inherently defective and will itself have to yield after trial to a more realistic facing of the central problems that beset all efforts to test skill in composition above the most rudimentary level. Yet even if—or when—the "writing sample" shows its faults in its initial forthcoming trial, we should not, in my judgment, be content simply to return to the old forms of the test with a feeling that the failure of this experiment fully justifies its predecessor.

The dissatisfaction with oblique and objective testing has its source in a genuine concern for the processes of education, a matter often lost sight of by those entirely committed to the subtleties of measurement. It is, moreover, supported by the frustrations and fears of thousands, even tens of thousands, of people for whom the Board tests now often seem to constitute a power of decision outrageously disproportionate to their actual value. If there is any reasonable and safe way of coming to terms with that concern for the processes of education and with the worries of countless parents and children, we ought to try it even at some cost to goals which testing experts have in large measure considered inviolate.

Permit me to preface my proposal with the tentative criticism that testers may have honored the relative where they should have honored the absolute and may have sacrificed much for the absolute where they might well have satisfied themselves and their customers with the relative. Let me compound the mystification in this synonym: they have been satisfied to accept what people do rather than what they should do but unwilling to let them be what they are.

I speak, of course, about language, not about morality, and the first part of my conundrum refers to the practice of accepting the common locution as a satisfactory one, a practice encouraged by the deliberations, if not always by the dogmas, of modern linguists. In my experience, most of the error—for such I think it to be—in this matter is one of omission rather than of commission. No test-maker has yet, so far as I know, accepted "ain't" or "aren't I?" as a correct response to a testing problem. Most have been wary even of allowing such disputed locutions as "due to" without a nominal referent or "hadn't ought" or "most perfect." But the lapses have nonetheless been condoned by the simple process of evasion.

The rationale for omitting items for which the common locution is neither elegant nor precise from a sophisticated test in the use of language, or ignoring the results if they get into a test by accident is, I am sure, statistically unassailable. If nearly all the examinees respond with a common but ineffective locution or, in some situations where the error is presented for correction, fail to respond at all, the item is deleted or not scored since, to use the statistician's phrase, it diminishes the validity of the score. Such items, in one analyst's terms, are "wasted."* Not for a minute would I suggest that there is among testers a conspiracy to establish the language of the uneducated as a norm or even to hasten the processes of linguistic change. All the testers I have known are—in speech, at least—literate persons and as fastidious as the English teacher about linguistic conventions. It is nonetheless true that the net result favors latitude rather than rigor, what *is* rather than what should be.

The present English Composition Test, for example, very rarely carries an item requiring correct use of the apostrophe, either in possessives or in contractions; it has entirely proscribed the distinctions between "that" and "which" in relative clauses and the distinctive punctuation of restrictive and nonrestrictive expressions. It ignores misspelling. In matters which rise above elementary conventions, like those of parallel construction, inversion, precision in the choice of adversatives, and the like, it is often forced by the statistical rationale to accept forms it can condone only by the bleak admission that rottenness is indigenous to the state of American usage.

Latitudinarian as it has come to be in its dogma, the test is often decried as unreasonably difficult when, that is, it is not being decried as unreasonably easy. The simple statistical fact

---

*** Analysis of ECT Data by ETS staff, Nov. 30, 1957.**

that the farther west of the Alleghenies you go, the lower the average score becomes may be a coefficient of the fact that the farther west you go, the smaller number of examinees you find taking the test, but I am inclined to doubt it. I suspect that, instead, we have here some ammunition for those who charge that these mass tests, even in their broad-church form, contain unconscious biases, that they are, more than we admit, instruments designed to play the tunes characteristic of the metropolitan East and therefore elicit only caterwauling or silence from those to whom their frets and valves are alien.

The matter is no joking one simply because it is as hard to establish the grounds of an absolute demand in language as it is to determine what a relative standard is relative to. "Where's it at?" may be a respectable locution in Indiana, but in Massachusetts it is unmistakably a vulgarism. The more subtle you try to make a test of language, the more you attend to diction and syntax, the more you find culturally validated differences; and the more differences you find, the more you are forced to widen the gauge or to eliminate the measurement altogether.

The problem, as I see it, is that testers are forced, by their determination to make one test do for all, to fight themselves. The more they attempt to measure the entire population on one scale, the more they must limit themselves in what they test. The result is bound to be as unsatisfactory to those who hold in reverence the linguistic conventions inherited from seventeenth- and eighteenth-century English plain style (the Northeast) as to those for whom the plain style is a corset (the South) and those for whom indigenous locutions have almost as much importance as those locutions mined from the mother lode (the Middle West and West).

If it is true that diction and syntax vary significantly from region to region in this country—the syntax of the Midwest and West being more laconic than that of the East, the diction of the South more periphrastic, openly metaphorical, and Latinate than that of East or West—then test-makers who want to deal with anything that might properly be called "elegance" in style face a choice between plumping for one to the disadvantage of those committed to another or devising some means for admitting the authority of all.

My preference is for the first choice—that is, provided that the one plumped for is the style which I find most pleasing. Judging from my reaction to the current enforced latitude on some items of the ECT, I doubt that I would cheer so heartily for the election of Southern-fried diction. Is the matter, then, purely a relative one? At the risk of rousing indignation, I must say that I think it is not. If the language we want to establish as the norm for educated people in the United States is that best suited to rational discourse, I believe that the language of the East has an edge on that of the other sections of the country. It may be less colorful; it may be less "natural," whatever that means; but it is more precise, more strictly ordered, more amenable to analytic procedures.

If I qualify this claim, as I do, by admitting that books of prose discourse written in the West, Midwest and South are similar in style to those written in the East, I shall have to compound my offense—or offensiveness—by observing that the style for such books is largely set in the East. New York City dominates American publishing, and the editors of the large publishing houses temper the manuscripts they read by eastern tuning-forks. But our concern is not really with the style of mature people who write books, only with that of young people who write examinations. And, in fact, we are not so much concerned with their free writing as with the kinds of problems to which they can respond on a test.

Short of creating a national academy, like that in France, for regulating official use of the language or, in the fashion of nineteenth-century England, a Society for the Purification of the American Language in the Tradition of the Eastern States, I see no feasible way of making my parochial view of stylistic truth prevail. In theory, it prevails today, of course, and that is exactly the trouble, as far as our tests are concerned. The expectation of a particular standard governs the construction of the tests, but the demand for performance consistent with that standard is modified and substantially weakened by the exigencies of statistical calculation.

In the face of this problem, I see no alternative, if there is to be testing of prose style in matters beyond basic conventions, to the second choice—that of devising some means of admitting the authority of the various national patterns of language. I may be wrong about this matter, but it seems to me that the problem is not solvable on a national test. My proposal therefore entails more than one test and more than one center of reading. And by "more than one center of reading" I mean to imply something other than a mere distribution of reading centers to which the law from a Princeton Sianai goes forth by mail or teletype.

What I should like to see tried is this: several tests of comparable difficulty but differing in content offered throughout the country, the administration to be followed at once by a large number of reading sessions held in the center of any region populous enough to furnish from five to ten thousand papers. For each a chief examiner, working with other examiners of the area, would set the answers and the standards to be applied. Atlanta Elsie's paper would thus fall into the hands of people for whom her manner is as familiar as breakfast hominy, and Boston Johnny's into those of readers who prefer their corn uncontaminated by lye or soda water.

A further benefit that might derive from such decentralization would affect the character of the test itself. In the first place, it would make possible the retention of many items now deleted, whether or not they produced a high incidence of response, simply because it would be reasonable to assume that the failure to respond indicated a failure to learn the accepted locutions of a region. The assumption, of course, would be—as it must be—that the teaching was consistent with the syntax and idiom of the regions, and localized inadequacies in performance would be a fairly sound indicator of poor instruction. The feedback from test to curriculum might then become more definite than it now is—a condition I believe to be eminently desirable despite the fact that testers are almost paranoid in their fear of being accused of dictating courses of study.

In the second place, it might become possible, given (let us say) thirty reading centers instead of one, to reintroduce some direct writing to the test in the form not of complete essays, but of paragraph responses to carefully contrived problems. Here, of course, we run into the familiar bugaboo of all essay testing—the known unreliability of essay scoring and the intrinsic unreliability of a single essay as evidence of a candidate's competence in writing. In some measure, both difficulties can be diminished by setting the candidate three or four problems rather than a single one in the testing period, a practice followed by the Board for a short time during World War II. Repeated attacks on writing problems, even within a brief period of testing time, offer a better indication of a student's compositional competence than a single extended essay. And it is my experience that single paragraphs can be read with greater reliability than a many-paragraphed essay. Even more important, however, is the opportunity a large number of reading centers offers for establishing among the readers a degree of agreement about standards, something manifestly impossible without rigorous and continued discipline if the number of readers in a group gets much higher than ten or fifteen.

The objections to such a plan for introducing greater rigor in a given reading of papers by allowing that group to define its standards are many. It would be more costly, of course; but I refuse to consider that a serious objection since it is entirely uneconomical to economize in a testing procedure on which the expenditure of an extra dollar or two for each candidate is insignificant in comparison with the total expenditure for education which the tests help to initiate. It would be administratively onerous; but if it is not administratively impossible, I see no objection in that: what are administrators for if not to carry burdens, each administrator a kind of *pons asinorum*, so to speak? Finally, it would create multiple standards of excellence instead of a single one. That, quite clearly, is the objection on which the entire proposal must stand or fall.

The sanctity of a single standard, as I suggested earlier, is more than a little sullied by the compromises that must be made to ensure it. For while no one today can claim that a score of 662 means one thing about a candidate from a rural Rocky Mountain school and another about a candidate from Phillips Exeter Academy, there is more than a suspicion that the score does not mean enough about either. What would happen if the 200-800 scale for one region were markedly different from the 200-800 scale in another? For ninety percent of the college-bound students in every region but the

Northeast, I think it would make no difference at all simply because that ninety percent will be attending colleges in the region of residence. And for that ninety percent the discrimination possible once the whole scale was confined to it should, I believe, be considerably more valuable to students and to college admissions officers than it now is. What, then, of the ten percent who go outside the region and of the diaspora of Northeasterners? I see no more objection to a formula of adjustment whereby the scores from one region are scaled down and of another raised, preliminary to comparison of candidates, than to the present uniform scale which, as a recent article in *Life* magazine shows, consigns institutions to categories of real or imagined quality on the basis of their median scores.

We are all sufficiently indoctrinated in this country with the scientific concepts of virtue to admire procedures that are simple and clean-cut and to suspect those that seem untidy and varying in their controls. But we dare not sacrifice all for tidiness and uniformity in a procedure which has a kind of magical power over curriculum and human hopes.

Last spring I talked with the admissions tutor of one of the colleges at Oxford, a distinguished literary scholar who manages to run the admissions job with just a little clerical help. His college selects one out of nine applicants. The examination papers, of course, are read by others, in a leisurely fashion, over the summer preceding the fall term, and the chosen applicants are notified of their election early in September. I asked the gentleman if that procedure were not hard on the eight candidates rejected out of every nine. He brushed my query aside with the remark that Americans believe in equality, Englishmen in liberty.

The comparison is scarcely a new one, and it occurred to me at the time that the liberty he spoke of seemed to be rather more useful to the college than to the candidate: one had the liberty to choose carefully and at leisure, the other the liberty to go hang if he were not chosen. But there is bite in the remark all the same. Equality we want, even in so small a matter as testing for competence in composition, but equality is a relative matter, not a universal one. We may legislate it in human affairs that bear on man's dignity and his opportunity for development; but it is entirely possible to legislate equality and produce inequity instead of justice. And, further, it is possible to be so committed to a uniform equality that justifiable local differences are penalized. That, I believe, is an inevitable consequence of any nation-wide test that tries to discriminate those who write well from those who write poorly, when "well" and "poorly" refer to an accomplishment greater than simple correctness in the use of isolated linguistic items.

WILMARTH H. STARR,
Foreign Language Testing Program, Modern Language Association of America

Competency First: New Tests in Foreign Languages

ANATOLE FRANCE is reported to have said that the world would come to an end from a glut of reading. I think that if he had been here this late October afternoon, he would have been sure that Armageddon was close at hand, but he would have revised his statement to read that the world would come to an end through a glut of testing. I assure you that I do not mean that this thought would put on the side of the angels your speaker who has just presided over the design and development of 62 tests of seven competencies, in five languages, and their administration to well over 30,000 cases, geographically distributed from Copenhagen to Guadalajara in the summer just past. But how did this come about?

I ask this question because it gives me the comfort of resorting to my professional habit of never getting into a subject without submitting an historical preamble. Yet it seems to me that it is exactly such an historical perspective that we must all keep before us when we would address ourselves to any question related to foreign languages. I am sure that you are all aware of the impact that World War II had upon the educational system—particularly that phase of impact which brought the realization to America that, whether we would have it so or not, our accumulated global responsibilities have placed us in a position where the ability to communicate with other peoples in their own tongue and with understanding of their values as seen from within has assumed a primary importance. I am sure, too, that I need not remind this audience that this sudden realization had a profound effect on the teaching of foreign languages in our country. I need no more than remind you that until World War II the typical pattern of foreign language education in our country was a

QUALIFICATIONS FOR TEACHERS OF MODERN FOREIGN LANGUAGES

(A modification of the 1955 MLA Statement "Qualifications for Secondary School Teachers of Modern Foreign Languages," this version is prepared as a workguide for Test Construction Committees.)

COMPETENCE	SUPERIOR	GOOD	MINIMAL
Listening Comprehension	Ability to follow closely and with ease all types of standard speech, such as rapid or group conversation and mechanically transmitted speech.	Ability to understand conversation of normal tempo, lectures, and news broadcasts.	Ability to get the sense of what an educated native says when he is making a special effort to be understood and when he is speaking on a general and familiar subject.
Speaking	Ability to speak fluently, approximating native speech in vocabulary, intonation, and pronunciation. Ability to exchange ideas and to be at ease in social situations.	Ability to talk with a native without making glaring mistakes, and with a command of vocabulary and syntax sufficient to express one's thoughts in conversation at normal speed with reasonably good pronunciation.	Ability to read aloud and to talk on prepared topics (e.g. for classroom situations) without obvious faltering, and to use the common expressions needed for getting around in the foreign country, speaking with a pronunciation understandable to a native.
Reading	Ability to read almost as easily as in English, material of considerable difficulty.	Ability to read with immediate comprehension prose and verse of average difficulty and mature content.	Ability to grasp directly (i.e. without translating) the meaning of simple, non-technical prose, except for an occasional word.
Writing	Ability to write on a variety of subjects with idiomatic naturalness, ease of expression, and some feeling for the style of the language.	Ability to write a simple "free composition" such as a letter, with clarity and correctness in vocabulary, idiom and syntax.	Ability to write correctly sentences or paragraphs such as would be developed orally for classroom situations and to write a simple description or message without glaring errors.

Applied Linguistics	The "good" level of competency with additional knowledge of descriptive, comparative, and historical linguistics.	The "minimal" level of competency with additional knowledge of the development and present characteristics of the language.	Ability to apply to language teaching an understanding of the differences in the sound system, forms, and structures of the foreign language and English.
Culture	An enlightened understanding of the foreign people and their culture, such as is achieved through personal contact, through travel and residence abroad, through study of systematic descriptions of the foreign culture, and through study of literature and the arts.	The "minimal" level of competency with first-hand knowledge of some literary masterpieces and acquaintance with the geography, history, art, social customs, and contemporary civilization of the foreign people.	An awareness of language as an essential element of culture and an understanding of the principal ways in which the foreign culture differs from our own.
Professional Preparation	A mastery of recognized teaching methods, evidence of breadth and depth of professional outlook, and the ability to experiment with and evaluate new methods and techniques.	"Minimal" level of competency plus knowledge of the use of specialized techniques, such as audio-visual aids, and of the relation of language teaching to other areas of the curriculum. Ability to evaluate the professional literature of foreign language teaching.	Knowledge of the present-day objectives of the teaching of foreign languages as communication and an understanding of the methods and techniques for attaining these objectives.

two-year sequence of a foreign language, generally taught by people who could not speak the language they were teaching. Furthermore, they taught the *modern* foreign language as they themselves had been taught—in the classical tradition of grammatical analysis and with emphasis upon translation from the foreign language to English. Unfortunately this is still the pattern in too many schools, or should I say in too *few* schools since less than fifty per cent of American public secondary schools even offer a foreign language.

But an important and exciting counter-fact has been the steady development of new attitudes toward the teaching of modern foreign languages in our country since the war. This "quiet revolution," as it has been often called among foreign language teachers, has been characterized by four new emphases which I would describe as follows:

1. A new ordering of and stress upon the language skills—i.e. listening comprehension, speaking, reading, and writing;
2. The inclusion of knowledgeability about descriptive linguistics (not a new discipline in itself, but one new to the general foreign language teacher) in its applied aspects. This concern for applied linguistics is based on the assumption that insights into the *nature* of language are of advantage to the teacher in that they enable him to predict a.) the difficulties the student who has already acquired the habit patterns of English speech will have in encountering a new sound system and b.) the difficulties this same student will have in contrasting the new sound system with the new graphic system of the foreign language he is learning;
3. The growing conviction that cultural insights into the mentality and behavioral patterns of the peoples whose language is being learned have equal importance with the language skills; that in dealing with such cultural insights, the contemporary aspect has primacy over the historical, behavior and attitudes have primacy over the more traditional cata-

loging of monuments and artifacts; values are of more immediate concern than discussion about the quaint and the picturesque;

4. The conviction that the teaching of foreign languages is a profession with a methodology that is subject to scientific analysis, that the methodology can be taught to the novice teacher and that it can be demonstrated to have beneficial results in terms of the objectives implied by the foregoing three arguments.

A second dimension of historical perspective must also be brought to bear upon the question before us. It is the general and growing dissatisfaction with the pattern of American education which tends to qualify teachers quantitatively under the assumption, which has been honestly held, I am sure, that hours of contact with subject matter in the classroom, and credits gained thereby, would insure the necessary qualification to teach. But this pattern is valid only if it rests upon a guarantee that the hours of coverage are adequate for mastery of the subject to be taught, and that the credits earned have a value that implies some degree of excellence. I would not here raise the question of whether this theory applies equally to all subjects, but I would point out the discrepancy between the theory and the reality as it applies to my own discipline. The nature of foreign language learning is such that it involves in the mastery of the listening, speaking, reading and writing objectives skill learning as differentiated from subject matter coverage. Since not all human beings are able to master skills at the same rate or to the same degree of excellence, it would seem to me, as it has to many others, that measurement in these areas should endeavor to designate the degree of competency acquired, and that such competency cannot be measured quantitatively but must of necessity be measured qualitatively.

It is important at this point to keep in mind that if the skills of reading and writing alone were involved, only new directions would have to be added to time-honored measuring instruments such as those familiar to you all under the label of Cooperative Tests, Regents Examinations, or College Board examinations in foreign languages. But you will remember that in appraising the changes that have been taking place in foreign language learning I implied that the direct heritage of the methods of teaching the classical languages, the grammar-translation approach, has been generally discredited in view of the fact that it did not produce people able to communicate with other peoples in their own language. Although the tests mentioned above have been somewhat affected by the new emphases in language learning as I have been describing them, they still have a tendency to emphasize lexical items as contrasted with manipulative skill, grammar knowledge as contrasted with automatic habit patterns, and translation abilities as contrasted with fluency in the language.

Even if we assume, as does seem to be the case, that the above mentioned testing instruments can be expected to undergo the changes implied by the new emphases on language learning, the two skills of listening comprehension and speaking are still inadequately accounted for and I must remind you again that in the new order of events these particular skills have assumed enormous proportions. I hasten to say that I am aware of, and that I welcome, the fact that the College Board examinations and the Cooperative Tests now have available optional listening comprehension tests for some languages, and that there is some indication that speaking tests may be in evidence before too long. People as familiar as you are, however, with test development are aware of the long time required for the development of new test forms through the above channels.

A third perspective of an historical nature now needs to be added to the above two. It is the extremely rapid, rather heterogeneous development of the application of electronic devices to language learning. Only the first few chapters of this particular story can be read to date, but the phenomenon known as the language laboratory (and at this point I cannot refrain from saying one that may be rendered obsolete nearly any day by the perfection of the teaching machine) has added a new dimension to language learning. Acquisition of audio-lingual skills, which in the past tended to be incidentally and casually taught, has been considerably refined by the potential that the language laboratory has for training, through systematic imitation and drill, those skills held to be so vital to modern foreign language learning.

Now the language laboratory, although it has provided us with new devices for testing and measurement, in the case of listening comprehension and speaking skills, has also added a new concern: that of the considerable expenses involved in the production and scoring of tapes and in their successful use as measurement devices for which large scale, reasonably sensitive, and desirably standardized electronic equipment is required. It is probable therefore that we could not have anticipated adequate batteries of tests in these two skills, nor indeed new tests in the others, that would be designed to measure competency and that would be tailored to meet the requirements of the new approaches to foreign language learning, had it not been for the possibility of tapping the monetary resources provided by the government funds and the wide scale human resources available from inside the profession once such funds were at hand.

May I assume at this point that I have made the following case: that the atmosphere surrounding language learning has notably changed in the last ten years or so; that a new emphasis has been stimulated in terms of competency measurement, and that the application of electronic devices to language learning have provided us with new instruments for training and thereby for measurement.

I should now like to describe to you two large scale testing programs which are in progress, both of which have as their declared purposes competency testing in strict accord with all the implications of the foregoing statements about the urgencies and emphases on foreign language learning. In 1955, as one of its many activities, the Steering Committee of the Foreign Language Program of the Modern Language Association of America published the Qualifications for Secondary School Teachers of Modern Foreign Languages—a statement which was subsequently endorsed by eighteen national and regional organizations of language teachers. The statement established three general levels of proficiency: minimal, good, and superior for seven areas of language teaching competencies 1) listening comprehension*, 2) speaking, 3) reading, 4) writing, 5) applied linguistics*, 6) culture, and 7) professional preparation.

In February of 1956, at a conference of national leaders representing the field of educational administration, the need for development of standardized proficiency tests as an aspect of teacher preparation and certification was endorsed. Obviously a statement of desiderata in relation to language teacher competencies, no matter how strongly representative of a consensus of the profession, would not be as effective as the situation demanded until standardized tests could be developed that would implement the description of competencies.

In the spring of 1959 the means to develop nationally standardized qualification tests for teachers of foreign languages were implemented under a contract between the United States Office of Education (NDEA—Title VI) and the Modern Language Association. The project, titled the MLA-Foreign Language Testing Program, is administered from the MLA Foreign Language Program Research Center at 70 Fifth Avenue, New York. From the beginning, the program has been dedicated to the principle that these tests would be produced by the profession and for the profession. Over 200 language teachers, most of whom are experienced in test design and development, have been involved in the production of preliminary test forms in the seven competencies and in the five languages, French, Spanish, Italian, Russian and German. It is these tests that were administered last summer to well over 2,000 teachers in pre- and post-test situations involving two forms of each test.

It is too early to predict the final format of the tests, but I submit that the preliminary versions are both imaginative and exciting. The listening comprehension and speaking tests are produced on tapes. Spoken by native speakers they deal with a variety of situations and dialogues which one could expect to encounter in the foreign country. The speaking tests involve reading aloud to test pronunciation, stress, accent, and intonation, a section which calls for imitation of a model, a section to test manipulative skills in terms of changing given structural patterns, and situations pictorially presented which involve manipulative skill and fluency in response. The reading tests are designed to measure sensitivity to style, shades of meaning, and comprehension of the message. The writing tests inculcate various devices such as reconstructing paragraphs, interlinear correction of error, completion of telegraphic messages, and structured compositions. The applied linguistics tests endeavor to measure the competency of the examinee both to handle principles of linguistics as a factor of his own skills, and to show his ability to adapt those principles to predicting areas of learning difficulty that are the inevitable results of the conflict between acquired habits in the native language and those to be acquired in the foreign language. The culture tests endeavor to measure competency in dealing with behavior, as well as interpreting literary and artistic production. The professional preparation tests aim at measuring the examinee's ability to control the methods most likely to produce competency in the six preceding areas.

---

* Originally designated as aural understanding and language analysis respectively.

A second large program is currently being initiated, also under Modern Language Association direction, which will test again in five languages the skill competencies of the language learner from grades 7 through 14.

Directors of the classroom testing project are Donald Walsh of the MLA Foreign Language Program Research Center and Professor Nelson Brooks of Yale. In announcing the contract, the United States Office of Education stated that "traditional methods of teaching foreign languages in the United States have emphasized the reading and writing skills. The new tests of speaking and hearing are expected to have a powerful effect on foreign language teaching since they will make it much easier for the teacher to emphasize all the four skills in keeping with the intent of the National Defense Education Act." Thus you will see that the above statement reflects what I have been saying previously about the realignment of emphasis toward the audio-lingual skills.

Together, the two sets of tests will thus presume to measure the competencies through all levels of language learning and teaching from grade seven upward. It is appropriate to mention to this audience that both of the test batteries are being designed, developed and administered in collaboration with

Educational Testing Service where Mrs. Miriam Bryan is now serving as the ETS Director of the MLA programs. It is too early of course to predict the results but it is the hope and indeed the conviction of all those concerned with the projects that these testing activities herald a new trend for testing in the subject matter disciplines, one that will be characterized by the production of nation-wide norms and standards of competence within the discipline throughout the whole spectrum of learning upward from the elementary grades and including the teacher.

An indication of general professional interest in the program is provided in the following quotation from the annual report of the 1959 meetings of the Council on Cooperation in Teacher Education: "The Council on Cooperation in Teacher Education recommends that research and development in proficiency examinations of all kinds be encouraged and subsidized. The Modern Language Association in proposing requirements and in developing tests for assessing competency in the teaching of foreign languages has set an example that may be followed by other academic and professional disciplines."

I am not one to predict the results of research and design previous to the time when the data is in and evaluated, but I am happy to offer this discussion to you as an example of how members of a subject matter discipline, when armed with the necessary resources, can exercise their responsibility to clarify professional standards by producing instruments to measure their own teaching competence, and those to measure the competencies their students acquire in the times in which we live.

APPENDIX A

Tests in the Modern Language Association Battery

---

1. Tests of Listening Comprehension

General Information: All stimulus material is recorded on tape. Response material is contained in test booklet. Answers are marked on standard IBM answer sheets. Each test has approximately the following composition:

Part A	Items 1-24	Four-choice response material (rejoinders) to statements heard on tape.
Part B	Items 24-45	Four-choice response material to dialogs heard on tape.
Part C	Items 46-55	Two-choice response material based on complex dialog heard on tape.

Forms developed:

				No. of Scores		
Language	Form	No. of Items	Answer Type	Reporting	Analysis	Time
French....	AF1	55	4cA & 2cA	1 (Total)	4	30-35
	BF1	55	4cA & 2cA	1 (Total)	4	30-35
German...	AG1	55	4cA & 2cA	1 (Total)	4	30-35
	BG1	55	4cA & 2cA	1 (Total)	4	30-35
Italian.....	AI1	56	4cA & 2cA	1 (Total)	4	30-35
	BI1	56	4cA & 2cA	1 (Total)	4	30-35
Russian....	AR1	55	4cA & 2cA	1 (Total)	4	30-35
	BR1	55	4cA & 2cA	1 (Total)	4	30-35
Spanish....	AS1	55	4cA & 2cA	1 (Total)	4	30-35
	BS1	55	4cA & 2cA	1 (Total)	4	30-35

2. Speaking Tests

General Information: Stimulus material is recorded on tape. In addition, stimulus material is contained in booklets provided candidates. Candidates in all languages use the same booklets. Response material is entirely recorded on tape. Each test has approximately the following composition:

Part	Content	Approx. No. of Scorable Units	Nature of Response
A	Repetition of sentences heard on tape.......	30	Structured
B	Reading of paragraph in foreign language...	30	Structured
C	Modification of sentences heard on tape.....	25	Structured
D	Description of pictures.....................	64	Free

Booklets provided candidates contain the reading selections and pictures to be described.

Forms Developed:

				No. of Scores		
Language	Form	Approx. Total Scorable Units	Answer Type	Reporting	Analysis	Time
French......	AF2	150	Free & Structured	1	7	35
	BF2	150	Free & Structured	1	7	35

Language	Form	Approx. Total Scorable Units	Answer Type	No. of Scores		Time
				Reporting	Analysis	
German...	AG2	150	Free & Structured	1	7	35
	BG2	150	Free & Structured	1	7	35
Italian.....	AI2	150	Free & Structured	1	7	35
	BI2	150	Free & Structured	1	7	35
Russian....	AR2	150	Free & Structured	1	7	35
	BR2	150	Free & Structured	1	7	35
Spanish....	AS2	150	Free & Structured	1	7	35
	BS2	150	Free & Structured	1	7	35

Candidate Booklets, Speaking Tests.

Form A-2
B-2

3. Reading Tests

General Information: All stimulus and response material is contained in test booklet. Answers are marked on standard IBM answer sheets. Each test has approximately the following composition:

Part	Content	No. of Items	Answer Type	Suggested Time
A	Sentence Completion	1-20	4cA	(60)
B	Connected Passages...	21-70	4cA	
C	Reading Speed........	71-100	2cA	(10)

Forms Developed:

Language	Form	Approx. No. of Items	Answer Types	No. of Scores		Time
				Reporting	Analysis	
French	AF3	100	4cA & 2cA	2 (1)	5	70
	BF3	100	4cA & 2cA	2	5	70
German	AG3	100	4cA & 2cA	2	5	70
	BG3	100	4cA & 2cA	2	5	70
Italian	AI3	100	4cA & 2cA	2	5	70
	BI3	100	4cA & 2cA	2	5	70
Russian	AR3	100	4cA & 2cA	2	5	70
	BR3	100	4cA & 2cA	2	5	70
Spanish	AS3	100	4cA & 2cA	2	5	70
	BS3	100	4cA & 2cA	2	5	70

(1) Reading Comprehension and Reading Speed.

4. Writing Tests:

General Information: All stimulus and response material is contained in test booklet. Answers are marked on standard IBM answer sheets and in the test booklet. Each test has approximately the following composition:

Section	Content	Spanish	French	Italian	German	Russian
I	Selection of best expression....	1-19	1-16	1-20	....	1-20
IIA	Incomplete context.......	20-49	17-46	21-50	1-60	21-50
IIB	Telegraphic.....	50-59	47-50	51-60	61-70	....
III	Interlinear......	60-70	50-96	61-90	71-115	51-80
IV	Essay...........	....	....	....	....	....

Answer Type		Suggested Times				
		Spanish	French	Italian	German	Russian
I	4cA	(10)	(10)	(10)		(10)
IIA	Structured	(25)	(20)	(25)	(20)	(15)
IIB	Structured				(15)	...
III	Structured	(25)	(30)	(25)	(20)	(25)
IV	Free	(30)	(30)	(30)	(35)	(40)

Forms Developed:

Language	Form	Approx. No. of Items	Answer Types	No. of Scores		Time
				Reporting	Analysis	
French	AF4	96 & Essay	4cA & 2cA	1	6	90
	BF4	96 & Essay	4cA & 2cA	1	6	90
German	AG4	115 & Essay	4cA & 2cA	1	6	90
	BG4	115 & Essay	4cA & 2cA	1	6	90
Italian	AI4	90 & Essay	4cA & 2cA	1	6	90
	BI4	90 & Essay	4cA & 2cA	1	6	90
Russian	AR4	80 & Essay	4cA & 2cA	1	6	90
	BR4	80 & Essay	4cA & 2cA	1	6	90
Spanish	AS4	79 & Essay	4cA & 2cA	1	6	90
	BS4	79 & Essay	4cA & 2cA	1	6	90

5. Applied Linguistics

General Information: All stimulus and response material is contained in the test booklet. Answers are marked on standard IBM answer sheets. Each test has 5-6 parts, depending on the language. The following composition is typical:

Part	Content
A	Phonetics
B	Morphology & Syntax
C	Writing System
D	General Linguistics
E	Historical & Comparative Linguistics

Forms Developed:

Language	Form	No. of Items	Answer Type	No. of Scores		Time
				Reporting	Analysis	
French	AF5	65	4cA	1	6	40
	BF5	65	4cA	1	6	40
German	AG5	65	4cA	1	6	40
	BG5	65	4cA	1	6	40
Italian	AI5	60	2cA-4cA	1	6	40
	BI5	60	2cA-4cA	1	6	40
Russian	AR5	65	4cA	1	6	40
	BR5	65	4cA	1	6	40
Spanish	AS5	65	4cA	1	6	40
	BS5	65	4cA	1	6	40

6. Civilization and Culture

General Information: All stimulus and response material is contained in test booklet. Answers are marked on standard IBM answer sheets. All content is spiralled. There are no separate sections or parts.

Forms Developed:

French	AF6	1-100	4cA	1	6	60
	BF6	1-100	4cA	1	6	60
German	AG6	1-100	4cA	1	6	60
	BG6	1-100	4cA	1	6	60
Italian	AI6	1-100	4cA	1	6	60
	BI6	1-100	4cA	1	6	60
Russian	AR6	1-99	4cA	1	6	60
	BR6	1-99	4cA	1	6	60
Spanish	AS6	1-100	4cA	1	6	60
	BS6	1-100	4cA	1	6	60

7. Professional Preparation

General Information: All stimulus and response material is contained in the test booklets. Answers are marked on standard IBM answer sheets. All content is spiralled. There are no separate sections or parts. Test material is in English and applies to all languages.

Forms Developed:

Form	No. of Items	Answer Type	No. of Scores Reporting	 Analysis	Time
A-7	1-85	4cA	1	1	70
B-7	1-85	4cA	1	1	70

IRVING LORGE,
Teachers College,
Columbia University

Estimating Structure in Prose

SINCE the turn of the century, schoolmen have been trying to obtain or to develop a valid, reliable and impartial method of appraising the comprehensibility of textbooks and related teaching materials. These attempts have considered the responses of pupils on examinations based on the texts they have read, on ratings and judgments of the understandability for children by teachers, librarians and professors, and on the expressed reactions of children to books. E. L. Thorndike's *Teachers Word Book* (1), published in 1921, provided a more objective basis for estimating the difficulty of texts in terms of relative frequency (or rarity) of the words used.

The concept of readability, however, received its primary impetus from the researches of Vogel and Washburne (2). They were asked to assign books for supplementary reading to appropriate grades. Instead of judging the difficulty of the books, they based the classification of each book on the tested reading ability of children who had read and *liked* the book. About 37,000 children, who had taken the paragraph-meaning section of the *Stanford Achievement Test*, indicated on a ballot each book read and liked in the year's supplementary reading.

Note: It is a pleasure for me to express my appreciation to the staff members of the Institute of Psychological Research, Teachers College, Columbia University, who made most of the tortuous analyses: Miss Mabel Wilcox, Mrs. Dorothy Heft, and Mrs. Luciana Visentini Steinzor; to my former graduate students, Dr. Joshua Fishman and Dr. Walter MacGinitie; to the College Entrance Examination Board and Educational Testing Service for supplying the candidates' General Compositions; to Dr. John Carroll for supplying the 150 passages from his Subtreasury with their factor scores; and to the hundreds of students in Freshman English in the School of General Studies, Columbia University, for their cooperation.

For each of the 700 titles which had received at least 25 favorable votes, the book's score was the median of the reading grade scores of the children who had read and liked the book. Thus, the Winnetka list (3) provided an empirically graded list of 700 supplementary books.

But what of new books? Or of books that had been read by fewer than 25 children? Vogel and Washburne tried to identify factors within the book that would predict its grade placement. They were the first to develop a multiple-regression formula predicting grade placement from four factors internal to the book, namely, the number of different words per thousand words, the number of uncommon words per thousand words, the number of simple sentences in 75 successive sentences, and the number of prepositions per thousand words. The Vogel and Washburne Grade Placement Formula (4) became the prototype for a series of multiple-regression equations which predicted a criterion from the internal aspects of the book—e.g., vocabulary load, sentence complexity, idea density and human interest.

No other internal element of a text added significantly to *estimated grade placement or difficulty*. In general, in most available multiple-regression equations, such as Lorge (5), Flesch (6), Dale-Chall (7), grade placement is estimated from the weighted contribution of at least two of the four elements. It should not be surprising, therefore, that grade placement estimated from vocabulary load, sentence complexity, idea density, and human interest of a text *is invariant for any rearrangement* of the sentences in a text. The Lorge Reading Grade Index for the Gettysburg Address, for example, would be the same no matter in what order the sentences are presented to the unwary reader.

No readability formula credits structure, or organization, or coherence, or sequence, or unity. While it may be assumed that nearly all published books have some structure or organization or unity, few teachers of English composition would grant that structure is an inevitable ingredient of student essays.

In the search for some valid and equitable basis for evaluating candidates' compositions and essays, the College Entrance

Examination Board and the Educational Testing Service, indeed, have devoted considerable research endeavor. For example, in May 1956, candidates for the CEEB Achievement Test in General Composition (8) were asked to write on the topic, "Pressures of Modern Society that Tend to Direct the Lives of Young People." As a basis for their essays, they were asked first to read a relatively easy passage of about 1900 words relevant to the assigned broad topic. The candidate was instructed to spend five minutes in giving the theme in one or two concise sentences, another 20 minutes in preparing an outline, and 75 minutes in writing and revising the essay. The candidates were informed that the essay was to be judged for five qualities: Mechanics, Style, Organization, Reasoning and Content.

Since two of the qualities to be considered in evaluating the candidate's compositions were related to structure, I requested the Educational Testing Service to send me a sample of the compositions. The purpose was to ascertain whether structure or style, as rated by the expert readers of the compositions, might be related to the Lorge Readability Indexes for the compositions. ETS sent the 194 compositions which had been used in estimating the reliabilities of the readers who had evaluated the General Composition. From these, a random sample of 69 compositions was selected. For each of these, a Lorge Readability Index was computed. The Lorge Readability Index is based on a count of the number of words written, the number of sentences, the number of prepositions, and the number of different hard words. These four counts are the basis for the values: average sentence length, the proportion of prepositional phrases, and the proportion of different hard words. The readers' ratings for each composition for Mechanics, Style, Organization, Reasoning and Content, as well as the candidate's Scholastic Aptitude Test Verbal and Mathematical aptitude scores, were supplied by ETS.

Since an IBM 650 was available, the complete matrix of intercorrelations among the 24 variables was computed, followed by a factor analysis and rotation by Kaiser's varimax (9) of the five qualities by two readers, the average sentence length, ratio of prepositional phrases, the ratio of hard words, and the SAT V and M scores.

The two qualities of major interest for the estimation of structure were the ratings for organization and style. According to the directions to the readers, "*Organization* had to do with the extent to which the writer has formulated a clear intention which lent order to his essay and had, as a consequence, arranged the blocks of material in a reasonable and effective way." Transition between paragraphs (or blocks of material) was considered under organization, but transition within paragraphs was considered a matter of style. (10) Style was to include "choice of words, construction of sentences, and the internal organization of paragraphs . . .," that is, ". . . internal construction of the transition of the paragraph had to do with transition between sentence elements and the general flow of language within a paragraph unit . . ."

The first orthogonal factor was identified as general academic ability; the second, as a specific for the first reader's ratings on organization, reasoning and content; and the third, a specific factor for the idea density (in terms of prepositional phrases) and word choice (in terms of different hard words) in the composition.

On the first orthogonal factor, academic ability, the loadings for Mechanics, Style, Organization, Reasoning and Content by the first reader were .80, .83, .78, .72 and .70 respectively; for the second reader, the loadings respectively were .83, .86, .81, .81 and .80. On the same first factor, the loading on the SAT Verbal score was .77 and on the SAT Mathematics score, .72. On the elements of the Lorge Readability Index, the loadings on the first factor were .04 for average sentence length, .34 for ratio of prepositional phrases and .32 for ratio of hard words.

Apparently, the readers cannot evaluate organization independently of the "halo" either of the general goodness of the object—i.e., the composition, or of the person—i.e., the candidate. The elements in the Lorge Readability Index, however, seem to be measuring aspects of the composition which are related more to the composition than to the person. Moreover, the Lorge Readability Index, as may have been expected on logical grounds, is related in a minor way to ratings for either organization or style. A sort of "haloed" general evaluative judgment is subsumed in the readers' ratings for the five qualities.

Another opportunity to attempt to evaluate structure and unity was afforded through John Carroll's massive and significant study of literary style. (11) Dr. Carroll collected a sample of 150 passages from various sources to represent a very wide range of content and style. Each of these passages was judged in terms of "semantic differential" adjectival scales and independently appraised in terms of objective measures. The average of the ratings given by eight expert judges for each composition on each of 29 semantic differential adjectival scales and the actual counts for each passage in terms of 39 objective measures—e.g., number of paragraphs, of sentences, of unmodified nouns, provided 68 scores for the matrix of intercorrelations.

On factor-analysis, the matrix was resolved as six orthogonal factors (or at least nearly orthogonal). These factors were identified as good-bad, personal-impersonal, ornamented-plain, abstract-concrete, serious-humorous, and descriptive-narrative. As must be recognized, this represents an important resolution of the 68 scores of literary performance into six independent components.

For my interest in appraising structure, attention was directed to three of the "semantic differential" adjectival scales: namely, those for rating the passage on a seven-point scale for meaningless-meaningful, hazy-clear, and chaotic-ordered. Of all the experts' judgments for each passage, these three seemed to me to be related most nearly to what intuitively may be considered structure.

These three adjectival judgment scales have loadings on Carroll's first factor, *good-bad,* of .61 for meaningful, of .63 for clear, and of .58 for ordered, but apparently are independent of any of the five other orthogonal factors identified by Carroll. None of the objective scores (which are similar to those used as elements in readability formulas) had any significant loading on the first factor. Apparently, expert judges using "semantic differential" adjectival scales utilized a sort of "general goodness" as an internalized referent for judging the passages for most of the 29 aspects.

It seemed useful to ascertain the extent to which the Carroll orthogonal factors may be related to the Lorge Readability Index. Dr. Carroll generously provided the 150 passages and the six factor scores for each passage. For each of these 150 passages, except the one that was an incomplete 300-word sentence from a will, the Lorge Readability Index was computed.

A Kaiser varimax rotation yielded at least two orthogonal factors: the first, defined by Carroll's *ornamented-plain* had loadings on average sentence length of .62, on ratio of prepositional phrases of .44, and on ratio of hard words of .41; suggesting that Carroll's ornamented-plain summarizes whatever elements are measured in the Lorge Readability Index. The second factor, identified by Carroll's *personal-impersonal* factor, is negatively related to Lorge's ratio of prepositional phrases with a loading of —.65 and the ratio of hard words with a loading of —.40. Apparently, structure or unity was not identified in Carroll's factors nor in the elements of the Lorge Readability Index.

Is structure or organization, then, so integral a part of a written passage that it cannot be evaluated or measured? Is it always subsumed in a halo of "general goodness" or general evaluative set? Perhaps structure or organization of a text is taken so much for granted that only exceptional deviations in organization are recognized. Some of you may remember your first difficulties with the structure of Dos Passos, or Gertrude Stein, or James Joyce. Perhaps some of you may remember Bartlett's (12) *Remembering*, in which he asked subjects to reproduce the Indian tale called the "War of the Ghosts." Undoubtedly, some of the memorial inadequacies in reproducing the original text are related to the Indian folk tale's unusual structure for British or American subjects who have internalized a "more logical" structure. Sequence, unity and coherence may be so well learned that experts accept most written material as structured. Of course handbooks on writing refer to the necessity for "good" structure in English prose. Textbook discussions about good structure imply that "structure," "coherence" and "good organization" are practically synonymous. Words such as "unity," "order," "lucidity," and "sequence," if not used synonymously, suggest at least a close association with the concept of "good" structure.

None of these terms, as yet, has been defined adequately. Some understanding of the range of concepts implied by structure can be had from the contexts in which the words are used.

Handbooks about writing usually do define "coherence," "organization," "unity," and "sequence." In defining *coherence*, for example, Greever and Jones (13) indicate that "coherence implies a close and natural sequence of parts, where every part of a sentence must have a clear and natural connection with the adjoining part. A sentence is coherent when one idea is completed at a time, and when the idea which naturally comes first in thought or sequence is placed first in the sentence." For Goodwin and Guill (14), coherence "is the natural and logical order of development of a sentence, paragraph of composition, through a single idea. To be coherent, words, phrases and clauses must stand as near as possible to the words they modify. The development of the main theme should follow a natural and logical order. Succeeding paragraphs should be an amplification of the topic paragraph." Baldwin (15) considers coherence to be synonymous with order, that is, "a composition is coherent when the people that listen to it follow it readily, and when each part prepares for the next. To be coherent, a composition must take hold of both subject matter and audience. It should be arranged according to a plan, since natural thought sequences are not orderly." Perry (16) indicates that coherence occurs when parts that belong together are placed together. Writing is incoherent when instead of saying all that one has to on one topic in one place, one writes a little, changes the subject, remembers something that might have been said on the first topic, says it, and after saying it, resumes the second topic. Perry emphasizes that coherence demands unity in the sense of orderly connection of thought rather than mere consecutiveness.

Organization for Burrows, Fereber, Jackson and Saunders (17) is "clarification from the verbalization of ideas," although they state that "the power of organization is not learned but follows 'need' in a concrete setting. It develops with the clearness of purpose for which specific material is being gathered and presented. Organization involves choosing relevent levels in terms of a specific problem. A feeling for organization is an innate part of everyone."

As defined by Greever and Jones, *unity* is "a combination of related ideas where each sentence contains only one thought." Baldwin calls unity "a principle of clearness involving expansion of an idea and the fixing of interest," and Goodwin and Guill state that "unity is present when a theme presents a complete whole and holds firmly to a point of view. To secure a well-organized whole, a definite plan must be used. The point of view should not shift within a sentence, and the beginning of a paragraph should bear strongly upon the central idea."

Burrows, Fereber, Jackson and Saunders indicate that *sequence* "is the proper 'bundling of material,'" and that "the ability to use sequential writing develops in children maturationally with the dawning of a 'sense of order.'"

These definitions, however, do not give the principle of structure nor suggestions for many procedures needed to achieve proper unity, coherence, or sequence in prose.

The basic problem, then, was to discover some operational demonstration of unity of prose. If unity means that sentences follow each other in an orderly fashion, then if any passage were to be fragmented into portions and put into random order, educated subjects should be able to reconstitute the portions into a useful order. If this random portion procedure were adopted, then two kinds of questions may be asked: (1) are people agreed as to the structuring of prose passages, and (2) if such agreement is found, is the reordering an aspect of unity, coherence, and structure?

For investigation of the operational approach, three stories from *Reader's Digest* of October 1954 and four passages from *Encore* of July 1944 were selected. Each of the three stories from *Reader's Digest* was divided into twelve or thirteen portions of approximately 150 words each. The passage entitled "Nature Man" was chosen because it presented several short themes (marsh birds, Captain Frank's life, Louisiana iris) loosely connected with each other, without any logical or chronological order. The passage on "World Bank" was selected as presenting two major themes (Eugene Black's efforts for the World Bank and the character of the bank itself) in a spiraling order—i.e., topic A leading to topic B, back to topic A, then again to topic B, etc. The piece called "Russian Story" was chosen because it presented one consecutive theme (Lt. Colonel Grigori Burlutski's activities in the Russian Army).

The four passages from *Encore* were each divided into their ten to thirteen constituent sentences. "Some Old Families" presented two major themes (description of Catamount, West Virginia and "Bill"); "The Assassination of President Garfield" presented several aspects of Garfield's life (freemasonry, scrapbook hobby, mastery of language and science); "On W. S. Gilbert" presented two major themes (people in Gilbert's serious plays versus those in Gilbert's fairyland); "The Last of R. S." presented one continuous theme (the last illness of R. S.).

The portions of each story and the sentences of each passage were randomized and numbered. A disarranged story of dissected portions, and a disarranged passage of sentences were collated in each of the possible combinations of two at a time. The sets were distributed as a homework assignment to students in freshman English classes (A1, A2) in the School of General Studies at Columbia University. Each student, thus, was given a set of scrambled portions which he was to rearrange into a complete story, and a set of scrambled sentences for arrangement into a paragraph. Students were asked to do the assignment independently, and to record what they judged to be the correct order by writing the numbers of the portions and of the sentences on a special answer sheet.

From the returned responses, a random sample was selected. For each of the three stories and four passages, the first 48 papers from the randomized pile were analyzed.

The objectives were: first, to ascertain whether these subjects would agree with one another in the ordering of the disarranged texts; and second, to ascertain how many breaks in sequence there would be. For estimating agreement in order, Kendall's coefficient of concordance was used. For estimating recognition of sequence, the number of breaks from published sequence was obtained.

Intuitively, for the passages with disarranged portions the expected order from most structured to least was expected to be: "Russian Story," "Nature Man" and "World Bank." The subjects agreed with the ranks given the disarranged portions in terms of Kendall's (18) coefficient of concordance as .70, .86 and .51, which put the passage considered least structured in the last position but did not put the other passages in the "expected" order. But, on the average number of departures from sequence, the expected was found—i.e., 2.4, 2.7 and 6.5, or in percentage terms about 16%, 20% and 43% departure from sequence.

For the disarranged sentences, the intuitively judged order for structure was expected to be: "Some Old Families," "The Last of R. S.," "President Garfield," and "W. S. Gilbert." The agreement in assigned ranks did not follow the "expected" order: the coefficients of concordance were .72, .37, .52, and .37. The average departures from sequence, however, were more nearly in line with expectations. The averages were 3.4, 5.6, 5.4 and 9.3, and in percentage terms 21%, 36%, 46%, and 62%. If these two leads are suggestive, some departure from sequence may be used as a measure of structure.

Dr. Joshua Fishman and I used still another approach for estimating structure. For the same disarranged portions of the stories, graduate students at Teachers College were asked to find for each portion the other portion that goes best with it either forward or backward. The results placed the three stories correctly in the intuitive order of structuredness—i.e., "Russian Story," "Nature Man" and "World Bank."

On the bases of these several explorations, it seems that structure can be evaluated independently of "general goodness" or "evaluative" estimations. The arrangement test which is so useful for establishing the structure in a set of pictures or cartoons seems to have utility in getting at the structure in prose. The departures from published sequence, or the departures from an order accepted by consensus may serve as the basis of a new approach for judging structure—although I doubt that it is an approach that will be relished by either teachers of composition or by expert readers.

In general, subjects found it easier to put portions into proper order than to rearrange sentences into proper sequence. This undoubtedly reflects the redundancy of more information given within a portion. Conversely, it is more difficult to put sentences into structural sequence because of the lack of enough redundancy.

The methods, though not new, may be applied to measure the connectedness of the text materials. Indeed, the measure may be an addition to the ratings by expert judgment of coherence and organization in written prose.

Perhaps a partial solution for the problems of rating or evaluating printed materials and candidates' compositions is to get rid of the excess in synonymy by restricting the evaluative words to two: (1) *coherence*, for sequence from sentence to sentence and from paragraph to paragraph, and (2) *organization* for clarity in the development and expansion of an idea.

REFERENCES

1. Thorndike, Edward L. *The teachers word book*. New York: Teachers College, Columbia University, 1921.
2. Vogel, Mabel and Washburne, Carleton. An objective method of determining grade placement of children's reading material. *The Elementary School Journal*, January 1938, 38, 373-381.
3. Washburne, Carleton and Vogel, Mabel. *Winnetka Graded Book List*. Chicago American Library Association, 1926.
4. Washburne, Carleton and Morphett, Mabel Vogel. Grade placement of children's books. *The Elementary School Journal*, January 1938, 38, 355-364.
5. Lorge, Irving. The Lorge Formula for estimating difficulty of reading materials. New York: Teachers College, Columbia University, Bureau of Publications, 1959.
6. Flesch, Rudolf F. Marks of readable style. New York: Teachers College, Columbia University, Bureau of Publications, 1943.
7. Dale, Edgar and Chall, Jeanne S. A formula for predicting readability. *Educational Research Bulletin*, January 1948, 27, 11-20.
8. Swineford, Frances. College Entrance Examination Board General Composition Test L. Princeton, N. J.: Educational Testing Service, 1956.
9. Kaiser, Henry F. Varimax solution for primary mental abilities. *Psychometrika*, June 1960, 25, 153-158.
10. Olsen, Marjorie. Summary of main findings on the validity of the 1955 College Board General Composition Test. Statistical report. Princeton, N. J.: Educational Testing Service, 1956.
11. Carroll, John B. Vectors of prose style, in Sebeok, Thomas A. (Ed.). Style in language. Boston: The Technology Press of MIT, and New York: John Wiley and Sons, 1960, 283-292.
12. Bartlett, Frederic Charles. *Remembering; a study in experimental and social psychology*. New York: Macmillan Co., 1932.

13. Greever, Garland and Jones, Easley S. *The Century handbook of writing*. New York: Century Co., 1921.
14. Goodwin, Mary Louise and Guill, K. G. *Students handbook of composition*. New York: Macmillan Co., 1918.
15. Baldwin, Charles Sears. *Composition oral and written*. New York: Longmans Green, 1922.
16. Perry, Frances M. *Progressive composition*. Yonkers, N. Y.: World Book, 1925.
17. Burrows, Alvina Treut; Fereber, June D.; Jackson, Doris C.; and Saunders, Dorothy Olton. *They all want to write*. Indianapolis, New York: Bobbs-Merrill, 1939.
18. Kendall, Maurice G. *Rank correlation methods*. London: C. Griffin, 1948.

PARTICIPANTS—1960 Invitational Conference on Testing Problems

ABRAHAM, A. A., Florida Agricultural and Mechanical University
ADAMS, Arthur S., American Council on Education
AGRE, Gene P., Educational Testing Service
ALBRIGHT, Frank S., West Orange (New Jersey) Public Schools
ALLEN, Frank, Jr., West Caldwell (New Jersey) High School
ALMAN, John E., Boston University
ALVERSON, Thomas M., New York State Department of Education
ALVI, Sabir Ali, Indiana University
ANASTASI, Anne, Fordham University
ANDERSON, David J., Pennsylvania State Department of Public Instruction
ANDERSON, G. Ernest, Jr., Newtonville (Massachusetts) Public Schools
ANDERSON, Rose G., New York City
ANDERSON, Roy N., North Carolina State College
ANDERSON, Scarvia B., Educational Testing Service
ANDREE, Robert G., Rich Township High School, Olympia Fields, Illinois
ANDREWS, T. G., University of Maryland
ARMSTRONG, Fred, United States Steel Foundation
ARONOW, Miriam S., Bureau of Educational Research, New York City Schools
ARSENIAN, Seth, Springfield College, Springfield, Massachusetts
ASH, Philip, Inland Steel Company
ASH, Philip L., Jr., The Provost Marshal General's School, Fort Gordon, Georgia
ASHCRAFT, Kenneth B., Colorado State Department of Education
ATKINS, William H., Yeshiva University
AUSTIN, David B., Teachers College, Columbia University
AZUMA, Hiroshi, Educational Testing Service

BAIER, Donald E., General Electric Company
BAKER, Frank B., University of Maryland
BALL, Albert T., Monmouth College, New Jersey
BANNON, Charles J., Waterbury (Connecticut) Department of Education
BARBARA, Sister, Fordham University
BARDACK, Herbert, New York State Department of Social Welfare
BARGMANN, Rolf E., Virginia Polytechnic Institute
BARKER, Roland, Montclair (New Jersey) Public Schools
BARRACANO, Albert, New York University
BARRONCINI, John, New York University
BATES, Margaret, Educational Testing Service
BEAIRD, James H., University of Nebraska
BECK, Saul, New York University
BELT, Sidney L., General Electric Company

BEMENT, Dorothy M., Northampton (Massachusetts) School for Girls
BENNETT, Clayton, North Central Association of Colleges and Secondary Schools, Chicago, Illinois
BENNETT, George K., The Psychological Corporation
BENNETT, Mrs. George K., Bronxville, New York
BENNETT, Ralph, New York City
BENSON, Arthur L., Educational Testing Service
BEREITER, Carl, Mellon Foundation, Vassar College
BERGER, Bernard, New York City Department of Personnel
BERGESEN, B. E., Jr., Personnel Press, Inc., Princeton, New Jersey
BIESHEUVEL, Simon, National Institute for Personnel Research, New York
BIXLER, Harold H., Western Carolina College, Cullowhee, North Carolina
BLEEKE, Donald E., Cranford (New Jersey) High School
BLETTER, Robert P., Columbia University
BLIGH, Harold F., World Book Company
BLITZ, Sidney, New York State Credit Union League
BLOMMERS, Paul, State University of Iowa
BLOOMER, Richard H., College of Education at Geneseo, New York
BOEHM, Charles H., Harrisburg (Pennsylvania) Department of Public Instruction
BOLAND, Ruth F., Cambridge (Massachusetts) School Department
BOLLENBACHER, Joan, Cincinnati (Ohio) Public Schools
BOROW, Henry, University of Minnesota
BOWMAN, Howard A., Los Angeles City Schools
BRACA, Susan, Oceanside (New York) Senior High School
BRANDT, Hyman, American Occupational Therapy Association, New York City
BREAM, Lois Gould, Cheltenham (Pennsylvania) High School
BRETNALL, William B., Educational Testing Service
BRIDGMAN, Donald S., Rowayton, Connecticut
BRINER, Lewis A., United Presbyterian Church Board of Christian Education, Philadelphia, Pennsylvania
BRISTOW, William H., Bureau of Curriculum Research, New York City
BRODINSKY, Ben, Arthur C. Croft Publications
BRODSKY, David J., Educational Testing Service
BROLYER, Cecil R., New York State Department of Civil Service
BROWN, David, Educational Testing Service
BROWN, F. Martin, Fountain Valley School, Colorado Springs, Colorado
BROWN, Fred S., Great Neck (New York) Public Schools
BROWN, Lee D., Science Research Associates
BROWN, W. F., University of Oklahoma
BRYAN, J. Ned, NCA Superior and Talented Student Project, Chicago, Illinois
BRYAN, Miriam M., Educational Testing Service
BUDDEKE, Sister Rita, Catholic University of America
BUHLER, Roald, Educational Testing Service

BURKE, James M., Norwalk (Connecticut) Public Schools
BURKE, Paul J., Bell Telephone Laboratories
BURNHAM, Paul S., Yale University
BUROS, Oscar K., Rutgers—The State University
BUSSIS, Anne, Educational Testing Service
BUSWELL, Guy T., American Educational Research Association, Washington, D. C.
BUYS, Janice C., Educational Testing Service

CAMPBELL, Cornelia, Educational Testing Service
CAMPBELL, Donald W., Newark (New Jersey) Public Schools
CAMPBELL, J. Chandler, Ohio State Department of Education
CAPPS, Marian P., Norfolk Division, Virginia State College
CARANCI, Joseph P., Cranston (Rhode Island) High School
CARITA, Sister M., Saint Elizabeth Convent, New Jersey
CARLSON, J. Spencer, University of Oregon
CARNEGIE, Elizabeth, Nursing Outlook, New York City
CARR, Mary B., Kingston (New York) Schools
CARROLL, John B., Harvard University
CARSTATER, Eugene D., Bureau of Naval Personnel
CARSTATER, Marie H., Falls Church (Virginia) High School
CHAPPELL, Bartlett E. S., New York Military Academy
CHAUNCEY, Henry, Educational Testing Service
CHAUSOW, Hymen M., Chicago Junior College
CHEW, William B., General Motors Corporation
CLARK, Albert H., American College of Life Underwriters, Philadelphia, Pennsylvania
CLARK, Philip I., California Test Bureau, Scarsdale, New York
CLEARY, J. Robert, Educational Testing Service
CLEMENT, William J., Bridgehampton (New York) High School
CLENDENEN, Dorothy M., The Psychological Corporation
CLIFF, Norman, Educational Testing Service
CLIFF, Rosemary, Educational Testing Service
COFER, Charles N., New York University
COFFMAN, William E., Educational Testing Service
CONNOLLY, John A., Educational Testing Service
CONRAD, Herbert S., U. S. Office of Education
CORY, Charles H., Philadelphia Personnel Department
COULSON, Roger W., Butler University, Indianapolis, Indiana
COUTANT, Madeleine F., Board of Cooperative Educational Services, Yorkville, New York
COX, Henry M., The University of Nebraska
COX, Thomas, Wallach-Cox Associates, Incorporated, New York City
CRONBACH, Lee J., Institute for Advanced Study, Princeton
CROOK, Robert B., Queens College
CROSSON, Wilhelmina, Palmer Memorial Institute, Sedalia, North Carolina
CUMMINGS, Mary B., Boston (Massachusetts) Public Schools

CUNLIFFE, Rex B., Rutgers University
CURRY, Robert P., Cincinnati (Ohio) Public Schools
CYNAMON, Manuel, Brooklyn College

DAILEY, John T., University of Pittsburgh
DALY, Alice T., New York State Education Department
DAMARIN, Fred L., Educational Testing Service
DAVIDSON, Helen H., The College of the City of New York
DAVIS, Anne D., Educational Testing Service
DAVIS, Ordie U., Kentucky State Department of Education
DAWE, J. Fred, Civil Service Commission of Canada
DEAN, E. Douglas, Princeton, New Jersey
DELKER, Harry A., South Dakota Department of Public Instruction
DE MOOSE, Norman M., Archdiocesan Vocational Service, New York City
DETCHEN, Lily, Chatham College
DICKSON, Gwen S., Kensington, Maryland
DIEDERICH, Paul B., Educational Testing Service
DIGGS, Franklin B., New York City Department of Personnel
DION, Robert, California Test Bureau
DODDS, Alice, Educational Testing Service
DOPPELT, Jerome E., The Psychological Corporation
DOWNES, Margaret, New York State Department of Civil Service
DOWNS, Harry S., University System of Georgia
DRAGOSITZ, Anna, Educational Testing Service
DRY, Raymond J., Life Insurance Agency Management Association, Hartford, Connecticut
DUBNICK, Lester, Huntington (Long Island) Board of Education
DUFF, Franklin L., World Book Company
DUGAN, Robert D., State Farm Insurance Company
DUKER, Sam, Brooklyn College
DUNN, Frances E., Brown University
DUNNE, Rev. William J., National Catholic Educational Association
DUROST, Walter N., World Book Company
DUTTON, Eugene, University of Alabama
DYER, Henry S., Educational Testing Service

EBEL, Robert L., Educational Testing Service
EDWARDS, Winifred E., Irvington (New Jersey) High School
EPSTEIN, Bertram, The College of the City of New York
EPSTEIN, Marion G., Educational Testing Service
ESSER, Barbara, Educational Testing Service

FAGIN, Barry, Marquette University
FAIRBANKS, Gar, Hofstra College
FAY, Paul J., New York State Department of Civil Service
FELDT, Leonard S., State University of Iowa
FENDRICK, Paul, Western Electric Company

FENSTERMACHER, Guy M., Educational Testing Service
FERGUSON, Leonard W., Life Insurance Agency Management Association
FIFER, Gordon, Hunter College
FINDLEY, Warren G., Georgia Board of Education
FINEGAN, Owen T., Gannon College
FINK, David, University of Maine
FISHELL, Kenneth N., University of Rochester
FLANAGAN, John C., American Institute for Research, Pittsburgh, Pennsylvania
FLEISCH, Sylvia, Boston University
FLESCH, Mary, Educational Testing Service
FLOYD, William, Columbia College (South Carolina)
FLYNT, Ralph C. M., U. S. Office of Education
FORMICA, Louis A., Norwalk (Connecticut) High School
FORRESTER, Gertrude, West Side High School, Newark, New Jersey
FORT, G. H., Atlanta (Georgia) Public Schools
FRANKFELDT, Eli, Research and Development Division, U. S. Army
FREDERIKSEN, Norman, Educational Testing Service
FREEMAN, Paul M., Educational Testing Service
FRENCH, Benjamin, New York Civil Service Department
FRENCH, John W., Educational Testing Service
FRIEDENBERG, Edgar Z., Brooklyn College
FRIEDMAN, Sidney, Bureau of Naval Personnel

GAGLIARDI, Ferdinand, Peekskill (New York) High School
GAINES, Lynn K., Educational Testing Service
GALLAGHER, Henrietta L., Educational Testing Service
GARD, Katherine, Educational Testing Service
GERBERICH, J. Raymond, University of Connecticut
GIAUDRONE, Angelo, Tacoma (Washington) Public Schools
GLASIER, Charles A., New York State Education Department
GLAUDE, Paul M., New York State Education Department
GODSHALK, Fred I., Educational Testing Service
GOLDMAN, Leo, Brooklyn College
GOLDSTEIN, Leo S., Cornell University Medical College
GOODE, Thelma, New York University
GORHAM, William A., American Institute for Research
GRAFF, Franklyn, Westport (Connecticut) Public Schools
GRANT, Donald L., American Telephone and Telegraph Company
GRAY, Mrs. Lyle Blaine, Remedial Reading Consultant, Baltimore, Maryland
GREEN, Robert E., National Research Council, Washington, D. C.
GREENE, Edward B., Chrysler Corporation
GULLIKSEN, Harold, Educational Testing Service

HAGIN, Rosa A., Irvington (New Jersey) Public Schools
HAGMAN, Elmer R., Greenwich (Connecticut) Board of Education

HALL, Alfred, Educational Testing Service
HALL, Robert G., Manter Hall School, Cambridge, Massachusetts
HALLOWELL, Elizabeth A., St. Francis College, Brooklyn, New York
HAMMOCK, Joseph, Bell Telephone Laboratories
HARPER, A. Edwin, Jr., Ecumenical Training Center, Stony Point, New York
HARTFORD, Donald L., IBM District 9 Education Center
HARTSHORNE, Nathaniel, Educational Testing Service
HARVEY, Philip R., Educational Testing Service
HASKELL, Paul G., Educational Testing Service
HASTINGS, J. Thomas, University of Illinois
HAUBRICH, Vernon, Hunter College
HAYWARD, Priscilla, Educational Testing Service
HEALY, Ernest A., Jr., Center for Psychological Service, Washington, D. C.
HEATON, Kenneth L., Heaton, Floyd and Watson, Philadelphia, Pennsylvania
HEIDGERD, Lloyd H., Educational Testing Service
HELEN LAWRENCE, Sister, Trinity College
HELM, Carl E., Educational Testing Service
HEMPHILL, John K., Educational Testing Service
HENRY, Sallyann, The Psychological Corporation
HERMAN, David O., The Psychological Corporation
HERRICK, C. James, Hartwick College, Oneonta, New York
HIERONYMUS, A. N., State University of Iowa
HILL, Carl M., Tennessee Agricultural and Industrial State University
HILLS, John R., University System of Georgia
HIMMLER, Merwin L., Pittsburgh (Pennsylvania) Public Schools
HINDSMAN, Ed, United States Office of Education
HOLLIS, Esther R., The Psychological Corporation
HOLMES, Fred E., Pennsylvania Department of Public Instruction
HOPKINS, Florence M., Educational Testing Service
HOPKINS, George W., South Carolina Department of Education
HOPKINS, Margaret L., Educational Testing Service
HOPMANN, Robert P., The Lutheran Church—Missouri Synod
HORROCKS, John E., Ohio State University
HORTON, Clark W., Dartmouth College
HOWARD, Allen H., University of Illinois
HOWARD, William E., Kent School, Kent, Connecticut
HOWELL, John, Educational Testing Service
HUDEN, John C., University of Vermont
HUGHES, John, International Business Machines Corporation
HUMPHRY, Betty J., Educational Testing Service
HUNT, Nile F., North Carolina Department of Public Instruction
HUNT, Thelma, George Washington University
HUNTER, Genevieve P., Fordham University
HUYSER, Robert J., Educational Testing Service
HYPPS, Irene C., Public Schools, Washington, D. C.

JACKA, Betty, Illinois State Scholarship Commission
JACKSON, Robert W. B., University of Toronto
JACOBS, Robert, U. S. Office of Education Services
JACOBS, Paul, Educational Testing Service
JANEBA, Hugo B., Rutherford (New Jersey) High School
JAROS, Andrew, University of Hartford
JASPEN, Nathan, New York University
JOHNSON, A. Pemberton, Newark College of Engineering
JORDAN, Arthur M., University of North Carolina
JUOLA, Arvo E., Michigan State University

KABACK, Goldie Ruth, The College of the City of New York
KAMMAN, James F., University of Illinois
KATZ, Martin, Educational Testing Service
KAVRUCK, Samuel, United States Office of Education
KELLER, Franklin J., Rutgers—The State University
KELLEY, H. Paul, University of Texas
KELLY, Thomas W., Educational Testing Service
KENNEY, Helen, Harvard University
KERSTING, Ethel, Educational Testing Service
KESSMAN, Maurice, Rochester Institute of Technology
KIMBALL, Elisabeth G., Educational Testing Service
KIRKWOOD, Virginia, Virginia State Department of Education
KLINE, William E., Board of Education of Baltimore County, Towson, Maryland
KLING, Frederick R., Educational Testing Service
KNERR, George, Pace College
KNOWLES, Edwin B., Pratt Institute
KOCH, John C., Jr., River Dell High School, Oradell, New Jersey
KOGAN, Leonard S., Community Service Society, New York City
KOGAN, Nathan, Educational Testing Service
KOUTSOPOULOS, C. J., Educational Testing Service
KUNOFSKY, Norma, New York State Department of Civil Service
KUNOFSKY, Solomon, New York State Department of Health
KURFMAN, Dana, Educational Testing Service
KVARACEUS, William C., Boston University

LADD, Carl E., Science Research Associates
LA DUKE, Charles V., United States Air Force
LAMBERT, Jane E., Educational Testing Service
LANG, Gerhard, Fairleigh-Dickinson University
LANGBERG, George, New York Board of Higher Education
LANKTON, Robert S., Detroit (Michigan) Public Schools
LANNHOLM, Gerald V., Educational Testing Service
LAPETINA, Francis, Peekskill (New York) High School
LARSON, Emilie G., Newton Centre (Massachusetts) High School
LECCESE, Victor, Oceanside (New York) High School
LENNON, Roger T., World Book Company

LEVINE, Harold, New York State Education Department
LEVINE, Milton, National Science Foundation
LINDQUIST, E. F., State University of Iowa
LINDSEY, John R., Department of Instructional Research, Detroit, Michigan
LITTERICK, William S., Keuka College, Keuka Park, New York
LIVINGSTON, W. D., Winthrop College, Rock Hill, South Carolina
LOMBARD, John, Educational Testing Service
LORD, Frederic M., Educational Testing Service
LORET, Peter G., Educational Testing Service
LORETAN, Joseph O., Board of Education, Brooklyn, New York
LORGE, Irving, Columbia University
LORGE, Sarah W., Board of Education, Brooklyn, New York
LOUGHRAN, George A., St. Francis College, Brooklyn, New York
LUCAS, Norma Lee, School District of Clayton, Missouri
LUISETTE, Joan, New York City Board of Education
LUNNEBORG, Clifford E., University of Texas
LYMAN, Howard B., University of Cincinnati
LYNCH, Eleanor, National League for Nursing, New York City

McCABE, William, New York University
McCALL, W. C., University of South Carolina
McCANN, Forbes E., McCann Associates, Philadelphia, Pennsylvania
McCOLLUM, Joyce, New York State Education Department
McCULLOUGH, C. William, Ohio State Department of Education
McCUSKER, Rev. William C., Regis High School, New York City
McDILL, Mrs. Thomas H., The Westminster Schools, Atlanta, Georgia
McGUIRE, Carson, The University of Texas
McINTIRE, Paul H., University of New Hampshire
McKENZIE, Francis W., Board of Education, Darien, Connecticut
McLAUGHLIN, Eugenia, New York State Civil Service Department
McLAUGHLIN, Kenneth F., U. S. Office of Education
McMANUS, Leo F., Jr., J. P. Cleaver Company, Worcester, Massachusetts
McQUEEN, Joan, Peekskill (New York) High School
McTERNAN, John W., State University Teachers College, Plattsburgh, New York
MACGINITIE, Walter H., Teachers College, Columbia University
MAIER, Milton, Educational Testing Service
MALCOLM, Donald, Educational Testing Service
MALKIEL, Judith, Educational Testing Service
MANGOLD, Rev. Eugene F., Jesuit Educational Association, New York City
MANLEY, Myron, Educational Testing Service
MANN, Sister Jacinta, Seton Hill College, Greensburg, Pennsylvania
MAROHN, Mildred, Lutheran Church—Missouri Synod
MARRON, Joseph E., United States Military Academy
MARTIN, Myra A., Ghana International School, Accra, Ghana
MARTIN, Harold C., Harvard University

MARTIN, H. T., Jr., University of Denver
MATHEWSON, Robert H., New York Board of Higher Education
MAXSON, Mildred H., Northern Valley Regional High School, Demarest, New Jersey
MAYNARD, Hull P., Loomis School, Windsor, Connecticut
MEDLEY, Donald M., Municipal College of New York City
MELVILLE, S. Donald, Educational Testing Service
MERENDA, Peter F., University of Rhode Island
MERNYK, Charlotte L., Brooklyn, New York
MERWIN, Jack C., University of Minnesota
MESSICK, Samuel J., Educational Testing Service
MILES, Mrs. J. R., Falls Church (Virginia) High School
MILHOLLAND, John E., University of Michigan
MILLARD, Kenneth A., Department of Army
MILLER, Harold T., Houghton Mifflin Company
MITCHELL, Blythe C., World Book Company
MITZEL, Harold E., Municipal College of New York City
MOLLENKOPF, William G., Procter and Gamble Company
MORRISON, Alexander W., Polytechnic Institute of Brooklyn
MUIR, Douglas, Department of Education, North Borneo
MUIRHEAD, Peter P., U. S. Office of Education
MULLEN, Toni, Educational Testing Service
MULLIGAN, Jeanne, New York University
MURRAY, John E., U. S. Naval Training Device Center, Port Washington, New York

NALLY, Thomas P., University of Rhode Island
NEVIN, Margaret, Educational Testing Service
NOISEUX, Ethel, New York State Department of Civil Service
NOLAN, David M., Educational Testing Service
NOLFO, Jacob, New York University
NOLL, Victor H., Michigan State University
NORMAN, Mary, University of Rochester
NORTH, Robert D., Educational Records Bureau
NOSOW, Sigmund, Michigan State University

OLSEN, Marjorie, Educational Testing Service
ORLEANS, Joseph B., George Washington High School, New York City
ORR, David B., American Institute for Research, Washington, D. C.
OTIS, C. Robert, California Test Bureau
OVERALL, John E., Johns Hopkins University

PACKARD, Albert G., Department of Education, Baltimore, Maryland
PAERATAKUL, Chawal, Indiana University
PALLONE, N. J., St. Francis College, Brooklyn, New York
PALMER, Orville, Educational Testing Service
PASHALION, Siroon, Herrold Associates, Flushing, New York
PAYETTE, Roland F., Educational Testing Service

PEMBERTON, W. A., University of Delaware
PERRY, William D., University of North Carolina
PETERKIN, A. Gordon, Westport (Connecticut) Public Schools
PFEIFFER, Ann, Baltimore (Maryland) Public Schools
PHILLIPS, Laura M., Roslyn, New York
PHILPOTT, Emily L., Queens College, Flushing, New York
PIERCE, Wendell H., Cincinnati (Ohio) Public Schools
PITCHER, Barbara, Educational Testing Service
PITT, A. Stuart, United States Naval Academy
POLLACK, Norman C., New York State Civil Service Department
POMEROY, Martha C., Indiana University
POPPENDIECK, Robert, United States Office of Education
POWERS, James, New York State Education Department
PRESCOTT, George A., Norwalk (Connecticut) Public Schools
PULLIAM, Rufus D., Georgia State Department of Education

QUINN, John S., Jr., World Book Company

RABINEAU, Louis, Pratt Institute
RABINOWITZ, William, Municipal College of New York City
RAPPARLIE, John H., The Owens-Illinois Glass Company
READ, Thomas, Maumee Valley (Ohio) Country Day School
REGAN, James J., U. S. Naval Training Device Center, Port Washington, New York
REID, John W., Newark College of Engineering
REINER, Harry, New York City Department of Personnel
REMMERS, Hermann H., Purdue University
RHUM, Gordon J., Iowa State Teachers College, Cedar Falls, Iowa
RICE, Edward W., North Plainfield (New Jersey) High School
RITENOUR, Louise R., Educational Testing Service
ROBERTS, S. O., Fisk University, Nashville, Tennessee
ROBLYER, William A., Psychological Services of Connecticut, New Haven, Connecticut
ROCA, Pablo, Pan American Union
ROEHRIG, Albert K., Bishop Hall, Andover, Massachusetts
ROSENGARTEN, William, Jr., Roslyn (New York) Public Schools
ROSENTHAL, Elsa, Educational Testing Service
ROSS, John, Educational Testing Service
RUBINFELD, William A., West Orange (New Jersey) High School
RULON, P. J., Harvard University
RUMMEL, J. Francis, University of Oregon
RUNKEL, Philip J., University of Illinois

SAIT, Edward, U. S. Army Enlisted Evaluation Center, Fort Harrison, Indianapolis, Indiana
SAMLER, Joseph, Veterans Administration, Washington, D. C.
SANDERS, Charles E., Louisville (Kentucky) Public Schools
SANFORD, Ruth, West Hemstead, Long Island

SANKIS, John, New York State Department of Civil Service
SAPAINCHIAY, Poj, Indiana University
SASAJIMA, Masu, New York Board of Higher Education
SATORN, Pinyo, Indiana University
SAUNDERS, D. R., Educational Testing Service
SCATES, Alice Y., United States Office of Education
SCHILD, Robert J., Nyack (New York) Public Schools
SCHRADER, W. B., Educational Testing Service
SCHUBERT, Arthur, New York University
SCOTT, C. Winfield, Rutgers—The State University
SCOTT, Winifred S., East Orange (New Jersey) Public Schools
SEASHORE, Harold, The Psychological Corporation
SEIBEL, Dean W., Educational Testing Service
SERLING, A. M., Educational Testing Service
SETZER, Charles J., New York City Department of Personnel
SFORZA, Richard F., New York State Department of Mental Hygiene
SHANNER, William M., California Test Bureau
SHARP, Catherine, Educational Testing Service
SHAYCOFT, Marion F., American Institute for Research, Washington, D. C.
SHEA, James E., New York State Department of Civil Service
SHIELDS, Mary, National League for Nursing, New York City
SHIELDS, O. L., Jefferson County Schools, Kentucky
SHIMBERG, Benjamin, Educational Testing Service
SIDWELL, Gray, Educational Testing Service
SILVERBERG, Edward B., New York City Department of Personnel
SILVEY, Herbert, Iowa State Teachers College
SIMPSON, Elizabeth A., Illinois Institute of Technology
SINGH, Anand, The New School for Social Research, New York City
SINGHAWISAI, Wilars, Indiana University
SINGLETON, Carlton M., World Book Company
SITGREAVES, Rosedith, Columbia University
SKINNER, James A., Corpus Christi (Texas) Independent School District
SMITH, Allan B., University of Connecticut
SMITH, Ann Z., Educational Testing Service
SMITH, D. D., Office of Naval Research
SMITH, Lyman J., Illinois State Scholarship Commission
SMITH, Marshal P., Trenton (New Jersey) State College
SMITH, Robert E., Educational Testing Service
SOLOMON, Robert J., Educational Testing Service
SOUTHER, Mary Taylor, Tower Hill School, Wilmington, Delaware
SPANEY, Emma, Queens College, Flushing, New York
SPAULDING, Geraldine, Educational Records Bureau
SPEER, George S., Illinois Institute of Technology
SPENCE, James R., New York State University
SPENCER, Richard E., The Pennsylvania State University
SPENCER, Wesley G., Cambridge, Massachusetts
SPRAGUE, Arthur R., Hunter College

SPREMULLI, Estelle, Educational Testing Service
STAHL, Suzanne, Educational Testing Service
STAKE, Robert, University of Nebraska
STANLEY, Julian C., University of Wisconsin
STARR, Wilmarth H., Modern Language Association, New York City
STECKLEIN, John E., University of Minnesota
STEVENS, Mary, West Hartford, Connecticut
STEWART, Mary, Medical Center, New York University
STEWART, Naomi S., Cranford, New Jersey
STICKELL, David W., Educational Testing Service
STILES, Grace Ellen, University of Rhode Island
STOKER, Howard, Florida State University
STONE, Paul T., Huntingdon College, Montgomery, Alabama
STRAUB, Ruth J., West Hartford (Connecticut) Public Schools
STREICHER, Samuel, Board of Examiners, Brooklyn, New York
STRICKER, Lawrence J., Educational Testing Service
STUART, William A., Educational Testing Service
SUPER, Donald E., Teachers College, Columbia University
SUSSMAN, Leah, Newark (New Jersey) Board of Education
SWANSON, Edward O., University of Minnesota
SWINEFORD, Frances, Educational Testing Service
SYMONDS, Mrs. Percival M., New York City

TANZER, Charles, LaSalle Junior High School, New York City
TAYLOR, Marvin, Queens College, Flushing, New York
THALHEIMER, Astrid H., Boston University
THOMAS, William F., University of Wisconsin
THOMPSON, Raymond, Educational Testing Service
THORNDIKE, Robert L., Teachers College, Columbia University
THORNE, Margaret A., Educational Testing Service
TITTLE, Carol J., Science Research Associates, Incorporated
TOOKER, Ellis D., Hartford (Connecticut) Board of Education
TRAIL, Stanley M., Rhode Island College
TRAVERS, Robert, University of Utah
TRAXLER, Arthur E., Educational Records Bureau
TREXLER, Laura M., Northern Valley Regional High School, Demarest, New Jersey
TRIGGS, Frances, Committee on Diagnostic Reading Test, Incorporated, Mountain Home, North Carolina
TROTH, William, Glen Ridge (New Jersey) High School
TULLY, G. Emerson, Florida State University
TURNBULL, William W., Educational Testing Service

UMANS, Shelley B., Brooklyn (New York) Board of Education
UPSHALL, Charles C., Eastman Kodak Company
URQUHART, Helen L., Brown University

VALLEY, John R., Educational Testing Service

VICKERY, K. N., Clemson College
VICKERY, Verna L., Mississippi State University
VITELES, Morris S., University of Pennsylvania

WAHLGREN, Hardy L., College of Education at Geneseo, New York
WALKER, Helen M., New York City
WALLACH, Philip C., Wallach-Cox Associates, Incorporated
WALLER, Raymond L., Allentown (Pennsylvania) School District
WALLMARK, Madeline M., Educational Testing Service
WALSH, John J., Boston College
WALTON, Wesley W., Educational Testing Service
WALTZMAN, Hal, New York State Department of Civil Service
WANTMAN, M. J., Educational Testing Service
WARD, Annie W., Volusia County (Florida) Schools
WARD, Lewis B., Harvard University
WATKINS, Richard W., Educational Testing Service
WATSON, Walter S., The Cooper Union
WEASNER, Harold, New York City
WEBB, Sam C., Emory University
WEINER, Max, Brooklyn College, Brooklyn, New York
WEINFELD, Frederic D., Science Research Associates
WEISS, Eleanor S., Educational Testing Service
WEITZ, Henry, Duke University
WEITZ, Joseph, New York University
WELLCK, A. A., University of New Mexico
WELNA, Cecilia, University of Hartford
WESMAN, Alexander G., The Psychological Corporation
WETTEROTH, William J., New York City Police Academy
WHITLA, Dean K., Harvard University
WIENER, Solomon, New York City Department of Personnel
WILKE, Marguerite M., Greenwich (Connecticut) Board of Education
WILKE, Walter H., New York University
WILLARD, Richard W., Massachusetts Institute of Technology
WILLEMIN, Louis P., U. S. Army Personnel Research
WILLEY, Clarence F., Norwich University, Northfield, Vermont
WILLEY, Lawrence V., Jr., Science Research Associates
WILLIAMS, Carla, U. S. Department of Health, Education and Welfare
WILLIAMS, Clarence M., University of Rochester
WILLINGHAM, Warren W., Georgia Institute of Technology
WILLIS, Leslie H., Bloomfield (New Jersey) Public Schools
WILSON, Phyllis C., Queens College, Flushing, New York
WINANS, S. David, New Jersey State Department of Education
WINTERBOTTOM, J. A., Educational Testing Service
WINTERS, John J., Jr., The New School for Social Research, New York City
WIRSIG, Jane, Educational Testing Service
WITHEY, Robert A., New Jersey State Department of Education
WOLMAN, Benjamin B., Queens College, Flushing, New York

WOMER, Frank B., University of Michigan
WRIGHT, Hugh King, The Lawrenceville School, Lawrenceville, New Jersey

ZALKIND, Sheldon S., New York University
ZUCKERMAN, Harold, New York City Board of Education